

Mixture Diffusion for Asset Pricing

Xin Liu*

First Version: July 15, 2016; Revised Version: December 6, 2016

Abstract

This paper proposes a general form of mixture diffusion process to model asset price dynamics, using a mixture of infinite number of parametric diffusions. We show that the underlying asset price dynamics of the risk-neutral distributions can be modeled precisely by the said mixture diffusions. Particularly, for mixture diffusion with random volatility, we can derive the explicit pricing formula for any path-dependent options that have a closed-form solution under Generalized Geometric Brownian Motion. A mixture diffusion based term structure model for instantaneous forward rate can price LIBOR Cap/Floor options consistently with the market.

Keywords: Local Volatility, Mixture Diffusion, Option, Risk-Neutral Distribution, Term Structure Model.

*DW Partners LP, 590 Madison Avenue, 13th Floor, New York, NY 10022; Xin.Liu@dwpartners.com.

This paper incorporates many important suggestions from Hongzhong Zhang's feedback. The author would like to thank Jean Wang, Ashish Misra for many valuable inputs to the manuscript.

1 Introduction

An Arrow-Debreu security is an elementary state-contingent claim that pays one dollar when a certain state occurs at a given future date and zero otherwise. In complete market, the price of an Arrow-Debreu security equals the risk-neutral probability density at the payoff state. [Breedon, Litzenberger(1978)] shows that the risk-neutral probability distribution implied by Arrow-Debreu prices can be extracted from European options if their prices exist across all possible strikes. Extensive research has been performed on how to extract the risk-neutral probability distribution from the option market because it provides valuable insight into the market risk preferences and expectations of asset valuations. [Jackwerth (2004)] and [Bahra (1997)] provide comprehensive surveys on this subject. In general, the problem of finding an asset price dynamics that is consistent with the implied risk-neutral distribution remains a challenging topic [Brigo (2002)]. This paper adapts the mixture model approach to find the said asset price dynamics.

Mixture model is an effective tool for analyzing financial time series data because it is flexible enough to capture the idiosyncrasies of financial data [McLachlan, Peel (2004)]. Particularly, the finite mixture models provide an exact solution for the asset price dynamics with intuitively appealing features by using weighted sums of Black-Scholes solutions. [Ritchey (1990)] suggests that the observed fat-tailed and skewed distributions can be modeled with finite normal mixture of independent Gaussian processes. [Alexander (2004)] introduces a parametrization of the normal mixture diffusion model that captures the short-term and long-term smile effect. [Gulisashvili (2012)] obtains the formula of mixing distributions for various diffusions with correlated or uncorrelated stochastic volatilities. [Brigo (2002)] and [Brigo *et al* (2002)] derive a stochastic differential equation (SDE) whose density evolves as a finite mixture of Gaussian densities.

This paper shows how to derive the diffusion process with an infinite number of mixture components, called mixture diffusion, such that its marginal density function fits precisely the risk-neutral distribution function across all maturities. There are three

types of risk-neutral mixture diffusions discussed in this paper: the mixture diffusion with deterministic volatility, random volatility and stochastic volatility. The mixture diffusion are shown to be mathematically workable and we can explicitly solve the SDE for mixture diffusion with random volatility. Furthermore, we can also derive the explicit pricing formula for any path-dependent options that have a closed-form solution with Generalized Geometric Brownian Motion (GGBM), while at the same time maintaining consistent European option prices with the market.

With mixture diffusion, we are able to combine the advantages of many existing asset pricing models into one single framework. First, mixture diffusion retains the simplicity of the Black-Scholes model and can price path-dependent options as the weighted average of option prices of underlying GGBM. Secondly, mixture diffusion has the flexibility of Dupire's local volatility model and we can calibrate the mixture diffusion such that its European option prices are consistent with the market across all maturities and strikes. Thirdly, mixture diffusion can capture the randomness of the volatility in the same manner as stochastic volatility models. Lastly, with the application for the term structure model of instantaneous forward rate, we demonstrate that the mixture diffusion is a very powerful tool in modeling asset dynamics. Specifically, the mixture diffusion term structure model is derived under a single risk-neutral measure, can price LIBOR Cap/Floor options that match exactly with the market prices and also provide consistent swaption prices.

The rest of paper is organized as follows: in Section 2, we introduce the concept of mixture diffusion with infinite mixture components and establish the theory on the existence of strong solution for the mixture diffusion with random or deterministic drift and volatility. In Section 3, we introduce the mixture of geometric diffusions and discuss how to re-parametrization them while keeping marginal density function the same. In Section 4, we present various parametrizations of geometric diffusions and derive the explicit solution for the mixing function. In Section 5, we derive the term structure

model of instantaneous forward rate that are consistent with the LIBOR market. Proofs of all results can be found in Appendix.

2 Asset Dynamics of Mixture Diffusions

2.1 Definition of Mixture Diffusion

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{Q})$ be a filtered probability space adapted to the standard Brownian motion $\{W_t : t \geq 0\}$. We explore various methods to find diffusion process in the probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{Q})$ such that its marginal density functions match the risk-neutral distributions. Because we concern only with fitting the risk-neutral distribution, we assume \mathbb{Q} is the same risk-neutral measure from which the risk-neutral probability density functions are derived. We assume the numeraire corresponding to the risk-neutral measure is the money market account $B(t)$, where $B(t) := e^{\int_0^t r(s) ds}$ and $r(t)$ is the risk-free interest rate. We will derive the risk-neutral asset dynamics directly under \mathbb{Q} and our approach does not involve the physical measure.

Let Θ be a Lebesgue measurable set in \mathbb{R}^n for some $n > 0$ and $T_0 > 0$ be a fixed future time. For every $\theta \in \Theta$, we assume that $\mu(x, t; \theta)$ and $\nu(x, t; \theta)$ are scalar functions such that the unique strong solution exists for the SDE

$$(2.1) \quad \begin{aligned} dX_t(\theta) &= \mu(X_t(\theta), t; \theta) dt + \sqrt{\nu(X_t(\theta), t; \theta)} dW_t \\ X_0(\theta) &= x_0 \end{aligned}$$

where $0 \leq t \leq T_0$. We denote the support interval of $X_t(\theta)$ as (a_0, ∞) where a_0 is typically either 0 or $-\infty$. We also assume the existence of the marginal probability density function of $X_t(\theta)$ with respect to the Lebesgue measure in (a_0, ∞) . We denote $P(\cdot, t; \theta)$ as the density function at time t with the parameter θ .

In the case of the finite mixture model, Θ is discrete, i.e., $\Theta = \{\theta_1, \dots, \theta_k\}$. The

mixture distribution based on $X_t(\theta)$ is defined as

$$\sum_{i=1}^k \lambda_i P(x, t; \theta_i),$$

where the mixture proportions $\{\lambda_1, \dots, \lambda_k\}$ satisfies $\lambda_i \geq 0$ and $\sum_{i=1}^k \lambda_i = 1$. Equivalently, we can write the mixture proportion in a functional form, denoted as $m(\theta)$,

$$m(\theta) = \sum_{i=1}^k \lambda_i \delta(\theta - \theta_i)$$

where $\delta(\cdot)$ is the Dirac Delta function. Then the mixture distribution becomes,

$$\int_{\Theta} m(\theta) P(x, t; \theta) d\theta$$

In this paper, we will use the functional form of mixture proportion as it is adaptive to infinite number of mixture components. We call $m(\cdot)$ the *Mixing Function* for the parametrized diffusion process (2.1) if it satisfies the following regularity conditions

C1. $m(\cdot)$ is integrable with

$$\int_{\Theta} m(\theta) d\theta = 1;$$

C2. For every $x \in (a_0, \infty)$ and $t \in (0, T_0]$, $m(\theta)P(x, t; \theta)$ is integrable with

$$\tilde{P}(x, t) := \int_{\Theta} m(\theta) P(x, t; \theta) d\theta > 0.$$

We call $\tilde{P}(x, t)$ the *Mixture Distribution* of the parametrized stochastic processes $X_t(\theta)$ with the mixing function $m(\theta)$. Our definition of mixing function ensures that $\tilde{P}(x, t)$ is a density function with the support interval (a_0, ∞) . We call a diffusion process $\{\tilde{X}_t : 0 \leq t \leq T_0\}$ the *Mixture Diffusion* of $\tilde{P}(x, t)$ if its marginal probability density function is the mixture distribution $\tilde{P}(\cdot, t)$ for every $0 < t \leq T_0$.

In the following sections, we will derive the explicit formula of the SDE for mixture diffusions and they are uniquely determined by the parametrized diffusion and the mixing

function. Therefore, we will also refer the parametrized diffusion (2.1) with the mixing function $m(\cdot)$ as *Mixture Diffusion* in the sense that it yields the diffusion process with the mixture distribution $\tilde{P}(x, t)$.

Since \mathbb{Q} is the risk-neutral measure, we can define the risk-neutral diffusion process under \mathbb{Q} as

$$(2.2) \quad \begin{aligned} \frac{dS_t}{S_t} &= r(t) dt + \sqrt{\beta(S_t, t)} dW_t \\ S_0 &= x_0 \end{aligned}$$

In this paper, we refer the mixture diffusion with the instantaneous drift term $r(t)x$ as *Risk-Neutral Mixture Diffusion*.

Even though a mixture diffusion can have the same marginal density functions as the risk-neutral density functions, in general, itself is not a risk-neutral process and we can not use it to value path-dependent options due to the existence of arbitrage opportunities. However, it is relative straightforward to derive the risk-neutral mixture diffusion with the same marginal density function from a general mixture diffusion. In this paper, we always first present the result with the general mixture diffusion and then derive its risk-neutral counterpart.

2.2 Mixture Diffusion with Random Drift and Volatility

The mixture diffusion can be naturally modeled with the diffusion process where the drift and volatility terms contain random variables. In fact, the mixture diffusion with random drift and volatility is a special case of diffusions with stochastic drift and volatility. To be more specific, we consider $(\tilde{X}_t, \tilde{V}_t)$, the diffusion with stochastic drift and volatility:

$$(2.3) \quad \begin{aligned} d\tilde{X}_t &= \mu(\tilde{X}_t, t; \tilde{V}_t) dt + \sqrt{\nu(\tilde{X}_t, t; \tilde{V}_t)} dW_t \\ d\tilde{V}_t &= 0 \end{aligned}$$

where it satisfies the initial condition $(\tilde{X}_0, \tilde{V}_0) = (x_0, V)$ and V is a n -dimensional random vector with the mixing function $m(\cdot)$ as its density function. It is clear that $\tilde{V}_t \equiv V$.

Consequently, (2.3) is reduced to the diffusion with random drift and volatility

$$(2.4) \quad d\tilde{X}_t = \mu(\tilde{X}_t, t; V) dt + \sqrt{\nu(\tilde{X}_t, t; V)} dW_t$$

Under suitable conditions (e.g., Assumption C_4' of Theorem 2.1), the conditional probability density function of \tilde{X}_t given V is $P(x, t; V)$. Therefore, the marginal probability density function of \tilde{X}_t is precisely the mixture distribution $\tilde{P}(x, t)$.

The random drift and volatility in (2.4) are random only at $t = 0$ and they are completely deterministic for $t > 0$. In contrast, the stochastic drift and volatility continue to evolve according to their diffusion equations as time passes. Though the stochastic volatility model is a more general approach to model diffusion process, however, it can be very difficult to find the closed-form solutions for its European option prices [Gulisashvili (2012), Hagan *et al* (2002)]. In contrast, mixture diffusion with random drift and volatility is easily mathematically workable: we can not only explicitly solve the SDE for the mixture of geometric diffusions (Section 3.1) but also have the closed-form solution for many path-dependent options (Theorem 2.5). More importantly, we can achieve these results while maintaining consistent European option prices with the market.

The following theorem provides the sufficient condition such that the diffusion with random drift and volatility admits a unique strong solution. This is the direct application of Theorem 2.2 of [Friedman (1975)] for the diffusion (2.3). For a vector $\theta = (\theta_1, \dots, \theta_n)$, we denote the Euclidean norm as $|\theta|^2 = \sum_{i=1}^n \theta_i^2$.

Theorem 2.1 *Let $\sigma(x, t; \theta) = \sqrt{\nu(x, t; \theta)}$. We assume*

C1. $|\mu(x, t; \theta)| + |\sigma(x, t; \theta)| \leq K(1 + |x| + |\theta|)$ for a positive constant K .

C2. For any $N > 0$, there exists a positive constant K_N such that

$$|\mu(x_1, t; \theta_1) - \mu(x_2, t; \theta_2)| + |\sigma(x_1, t; \theta_1) - \sigma(x_2, t; \theta_2)| \leq K_N(|x_1 - x_2| + |\theta_1 - \theta_2|)$$

for every $|x_1|, |x_2|, |\theta_1|, |\theta_2| \leq N$ and $0 \leq t \leq T_0$.

C3. Let $m : \Theta \mapsto [0, \infty)$ be a density function. We assume $\int_{\Theta} |\theta|^2 m(\theta) d\theta < \infty$.

C4. Let V be a n -dimensional random vector with the density function $m(\cdot)$. We assume V is adapted to \mathcal{F}_0 and is independent of $\mathcal{F}(W_t, 0 \leq t \leq T_0)$.

Then the diffusion process

$$(2.5) \quad \begin{aligned} d\tilde{X}_t &= \mu(\tilde{X}_t, t; V) dt + \sqrt{\nu(\tilde{X}_t, t; V)} dW_t \\ \tilde{X}_0 &= x_0 \end{aligned}$$

admits a unique strong solution with the marginal density function $\tilde{P}(x, t)$.

The risk-neutral counterpart of the mixture diffusion (2.5) can be derived using Fokker-Planck equation.

Proposition 2.2 *We assume the support interval of $X_t(\theta)$ is $(0, \infty)$ and the non-negative function $\beta(\cdot) : (0, \infty) \times [0, T_0] \times \Theta \mapsto [0, \infty)$ is a well-defined*

$$(2.6) \quad \beta(x, t; \theta) = \frac{1}{x^2} \left(\nu(x, t; \theta) + \frac{2}{P(x, t; \theta)} \int_0^x P(y, t; \theta) (r(t)y - \mu(y, t; \theta)) dy \right)$$

Assume the conditions in Theorem 2.1 hold, then the diffusion process

$$(2.7) \quad \begin{aligned} \frac{d\tilde{X}_t}{\tilde{X}_t} &= r(t) dt + \sqrt{\beta(\tilde{X}_t, t; V)} dW_t \\ \tilde{X}_0 &= x_0 \end{aligned}$$

admits a unique strong solution with the marginal density function $\tilde{P}(x, t)$.

From (2.7), it is clear that $\mathbb{E} \left(e^{-\int_0^t r(s) ds} \tilde{X}_t | V \right) = x_0$. Because Proposition 2.2 implies that the conditional probability distribution \tilde{X}_t given V is the same as that of $X_t(\theta)$ with $\theta = V$, we have

$$(2.8) \quad \mathbb{E} X_t(\theta) = B(t)x_0$$

for every $\theta \in \Theta$. The following corollary shows that the constraint (2.8) is also the result of the linear growth condition on $\beta(x, t; \theta)$ and $\nu(x, t; \theta)$.

Corollary 2.3 Assume $\beta(x, t; \theta)$ and $\nu(x, t; \theta)$ in (2.6) satisfy the linear growth condition,

$$(2.9) \quad x^2 \beta(x, t; \theta) + \nu(x, t; \theta) \leq K(\theta)(1 + x^2)$$

for some function $K(\cdot) : \Theta \mapsto \mathbb{R}^+$. Then we have for every $\theta \in \Theta$,

$$(2.10) \quad \mathbb{E}\mu(X_t(\theta), t; \theta) = r(t)\mathbb{E}X_t(\theta)$$

and consequently, the equation (2.8) holds true as well.

Later in this paper, we will parametrize the mixture diffusion as Generalized Geometric Brownian Motion (GGBM), i.e.,

$$(2.11) \quad \frac{dX_t(\theta)}{X_t(\theta)} = \mu(\theta, t) dt + \sigma(\theta, t) dW_t$$

where $\mu(\theta, t)$ and $\sigma(\theta, t)$ are deterministic scalar functions. The condition C1 of Theorem 2.1 requires that $\nu(\theta, t)$ and $\mu(\theta, t)$ are uniformly bounded. Unfortunately, most of our applications use unbounded parametrization and fail this requirement. Below we introduce a theorem designed specifically for unbounded parametrization. Though the theorem is stated in one-dimensional form, the same result also holds true when \tilde{X}_t and W_t are vectors if we interpret $|\cdot|$ as the Euclidean norm.

Theorem 2.4 Let $\sigma(x, t; \theta) = \sqrt{\nu(x, t; \theta)}$. We assume

C1. There exists a measurable function $f : \Theta \mapsto [0, \infty)$ such that

$$\begin{aligned} |\mu(x, t; \theta)| + |\sigma(x, t; \theta)| &\leq f(\theta)(1 + |x|) \\ |\mu(x_1, t; \theta) - \mu(x_2, t; \theta)| + |\sigma(x_1, t; \theta) - \sigma(x_2, t; \theta)| &\leq f(\theta)|x_1 - x_2| \end{aligned}$$

C2. Let $m : \Theta \mapsto [0, \infty)$ be a density function such that

$$(2.12) \quad \int_{\Theta} e^{C_0 f^2(\theta)} m(\theta) d\theta < \infty$$

where $C_0 = 10(T_0^2 + T_0)$.

C3. Let V be a n -dimensional random vector with the density function $m(\cdot)$. We assume V is adapted to \mathcal{F}_0 and is independent of $\mathcal{F}(W_t, 0 \leq t \leq T_0)$.

Then the diffusion process

$$(2.13) \quad \begin{aligned} d\tilde{X}_t &= \mu(\tilde{X}_t, t; V) dt + \sqrt{\nu(\tilde{X}_t, t; V)} dW_t \\ \tilde{X}_0 &= x_0 \end{aligned}$$

admits a unique strong solution with the marginal density function $\tilde{P}(x, t)$.

Then for the mixture of geometric diffusion, we can verify the integrability condition C3 directly by setting the dominant function $f(\theta)$ as

$$f(\theta) = \sup_{0 \leq t \leq T_0} |\mu(\theta, t)| + |\sigma(\theta, t)|.$$

The mixture diffusion with random drift and volatility has an important property in option valuation: because the random variable is known at $t = 0$, the expected value of the contingent claim conditional on the observed random variable is the just that of the contingent claim based on the underlying diffusion process of the given random value. Therefore, the expected value of a contingent claim is the weighted average of the expected values across all parametrized diffusions. That is the following theorem.

Theorem 2.5 *Assume $C(\tilde{X}.)$ is a contingent claim based on the mixture diffusion (2.13) and it has a finite expected value. Denote $C(X.(θ))$ as the contingent claim based on the parametrized diffusion process $X_t(θ)$ with the same payoff function. Then*

$$(2.14) \quad \mathbb{E}[C(\tilde{X}.)] = \int_{\Theta} m(\theta) \mathbb{E}[C(X.(θ))] d\theta$$

In Section 3, we use mixture of GGBM to derive the mixture diffusion such that it has the same marginal density function as the risk-neutral distribution. For many path-dependent options such as Asian options and Barrier options, we know their closed-form solutions under GGBM. With Theorem 2.5, we can obtain the explicit pricing formula

of these path-dependent options, while at the same time keep the values of European options consistent with the market.

As one example, we use Theorem 2.5 to value Variance Swap. The payoff of the variance swap for asset S_t at the expiry T is defined as

$$P(T, \sigma_{strike}) := \frac{1}{T} \int_0^T \left(\frac{dS_t}{S_t} \right)^2 dt - \sigma_{strike}^2$$

We assume the underlying asset follows the mixture diffusion (2.13) parametrized with GGBM (2.11), i.e.,

$$\frac{dS_t}{S_t} = \mu(V, t) dt + \sigma(V, t) dW_t.$$

Then the payoff function becomes

$$P(T, \sigma_{strike}) = \frac{1}{T} \int_0^T \sigma^2(V, t) dt - \sigma_{strike}^2.$$

Therefore, the expected value of the variance swap is,

$$(2.15) \quad \mathbb{E}P(T, \sigma_{strike}) = \int_{\Theta} m(\theta) \bar{\sigma}^2(\theta, T) d\theta - \sigma_{strike}^2.$$

where $\bar{\sigma}^2(\theta, t) := \frac{1}{t} \int_0^t \sigma^2(\theta, s) ds$.

2.3 Mixture Diffusion with Deterministic Drift and Volatility

The mixture diffusion with random drift and volatility requires that the mixing function is non-negative. In this section, we will derive the mixture diffusion whose mixing function can have negative values. The drift and volatility terms are deterministic in this case.

[Brigo (2002)] derives the explicit formula of the deterministic drift and volatility for finite mixture diffusions. In the theorem below, we present a similar result, but with the mixture of infinite diffusions. Note that the mixing function in the following theorem can have negative values or be time-dependent. Though all our results requires that mixing function to be time-independent, it is still important to understand how time-dependent

mixing function impacts on the mixture diffusion. The following theorem shows that the time-dependent mixing functions contributes an additional drift term, i.e., $\tilde{\alpha}(\cdot)$, to the SDE. This term disappears when the mixing function is time-independent.

Theorem 2.6 *Assume that the parametrized diffusion processes (2.1) and the time-dependent mixing function $m_t(\cdot)$ satisfies the conditions A1-A6 specified in Appendix.*

Define diffusion process \tilde{X}_t as

$$(2.16) \quad \begin{aligned} d\tilde{X}_t &= [\tilde{\mu}(\tilde{X}_t, t) + \tilde{\alpha}(\tilde{X}_t, t)] dt + \sqrt{\tilde{\nu}(\tilde{X}_t, t)} dW_t \\ \tilde{X}_0 &= x_0 \end{aligned}$$

where

$$\begin{aligned} \tilde{P}(x, t) &= \int_{\Theta} m_t(\theta) P(x, t; \theta) d\theta \\ \tilde{\alpha}(x, t) &= \frac{1}{\tilde{P}(x, t)} \int_{\Theta} \int_x^{\infty} \frac{\partial}{\partial t} m_t(\theta) P(y, t; \theta) dy d\theta \\ \tilde{\mu}(x, t) &= \frac{1}{\tilde{P}(x, t)} \int_{\Theta} m_t(\theta) P(x, t; \theta) \mu(x, t; \theta) d\theta \\ \tilde{\nu}(x, t) &= \frac{1}{\tilde{P}(x, t)} \int_{\Theta} m_t(\theta) P(x, t; \theta) \nu(x, t; \theta) d\theta \end{aligned}$$

Then SDE (2.16) admits a unique strong solution with the marginal probability density function $\tilde{P}(x, t)$.

The regularity conditions of Theorem 2.6 are such conditions to ensure all integrals and derivations in Theorem 2.6 are valid and the unique strong solution exists. Since the focus of this paper is to model risk-neutral diffusions, which typically involves some commonly used smooth functions, these regularity conditions are generally satisfied for our purposes.

The risk-neutral counterpart of the mixture diffusion (2.16) is presented below.

Proposition 2.7 *Assume the support interval of the mixture diffusion of Theorem 2.6 is $(0, \infty)$. We assume the non-negative function $\beta(\cdot) : (0, \infty) \times [0, T_0] \mapsto [0, \infty)$ is well-defined*

$$(2.17) \quad \beta(x, t) = \frac{1}{x^2} \left(\tilde{\nu}(x, t) + \frac{2}{\tilde{P}(x, t)} \int_0^x \tilde{P}(y, t) (r(t)y - \tilde{\mu}(y, t) - \tilde{\alpha}(y, t)) dy \right).$$

Under the assumption of the existence of an unique strong solution, the risk-neutral diffusion

$$(2.18) \quad \begin{aligned} \frac{d\tilde{X}_t}{\tilde{X}_t} &= r(t) dt + \sqrt{\beta(\tilde{X}_t, t)} dW_t \\ \tilde{X}_0 &= x_0 \end{aligned}$$

has the marginal density function $\tilde{P}(x, t)$ for every $t \in (0, T_0]$.

Similar to Corollary 2.3, the linear growth condition on $\beta(x, t)$ and $\tilde{\nu}(x, t)$ implies the following constraint on the expected value of the diffusion $X_t(\theta)$:

$$(2.19) \quad \int_{\Theta} m_t(\theta) \mathbb{E} X_t(\theta) d\theta = B(t)x_0$$

Corollary 2.8 Assume $\beta(x, t)$ and $\tilde{\nu}(x, t)$ in (2.17) satisfy the linear growth condition,

$$(2.20) \quad x^2 \beta(x, t) + \tilde{\nu}(x, t) \leq K(1 + x^2)$$

for a positive constant K . Assume that $\tilde{P}(\cdot, t)$ has a finite second moment, i.e.,

$$\int_{a_0}^{\infty} x^2 \tilde{P}(x, t) dx < \infty.$$

Then the equation (2.19) holds true.

Under the assumption of zero interest rate and no dividend, [Dupire (1994)] explicitly solves the instantaneous variance in the risk-neutral diffusion process when European option prices exist across all strikes and maturities. In this case, the risk-neutral mixture diffusion is identical to Dupire's local volatility. However, the mixture diffusion is a much more general approach because it has explicit parametrization of the underlying assets and does not require the existence of entire implied volatility surface to calibrate the model.

3 Mixture of Geometric Diffusions

3.1 Properties of Mixture of Geometric Diffusions

A simple and effective way to model the mixture diffusion is to parametrize the underlying diffusions as GGBM, i.e.,

$$(3.1) \quad \frac{dX_t(\theta)}{X_t(\theta)} = \mu(\theta, t) dt + \sqrt{\nu(\theta, t)} dW_t$$

where $\mu(\theta, t)$ and $\nu(\theta, t)$ are deterministic scalar functions of (θ, t) . We call the parametrization (3.1) as *Mixture of Geometric Diffusions*. It is commonly known that $X_t(\theta)$ has the lognormal probability density function, denoted as

$$L_t(x; \theta) := \frac{1}{x\sqrt{2\pi\nu(\theta, t)}} \exp\left(-\frac{(\log(x/F(t)) - u(\theta, t) + v(\theta, t)/2)^2}{2\nu(\theta, t)}\right)$$

where $F(t) = x_0 B(t)$ is the forward asset price and

$$(3.2) \quad u(\theta, t) := \int_0^t \mu(\theta, s) - r(s) ds, \quad v(\theta, t) := \int_0^t \nu(\theta, s) ds$$

It is often more convenient to model the *Moneyness* of Asset Price, defined as

$$Y_t(\theta) := \log(X_t(\theta)/F(t)).$$

Itô's Lemma shows that $Y_t(\theta)$ satisfies the SDE

$$\begin{aligned} dY_t(\theta) &= (\mu(\theta, t) - r(t) - \nu(\theta, t)/2) dt + \sqrt{\nu(\theta, t)} dW_t \\ Y_0(\theta) &= 0 \end{aligned}$$

$Y_t(\theta)$ then has the normal probability density function, denoted as

$$N_t(x; \theta) := \frac{1}{\sqrt{2\pi\nu(\theta, t)}} \exp\left(-\frac{(x - u(\theta, t) + v(\theta, t)/2)^2}{2\nu(\theta, t)}\right).$$

Assume \tilde{X}_t is the mixture diffusion with the lognormal distribution (such as derived in Theorem 2.1, 2.4 or 2.6 with the time-independent mixing function)

$$\tilde{L}_t(x) := \int_{\Theta} m(\theta) L_t(x; \theta) d\theta$$

and \tilde{Y}_t is the mixture diffusion with the normal distribution (such as derived in Theorem 2.1, 2.4 or 2.6 with the time-independent mixing function)

$$\tilde{N}_t(x) := \int_{\Theta} m(\theta) N_t(x; \theta) d\theta$$

It is straightforward to verify that both \tilde{Y}_t and $\log(\tilde{X}_t/F(t))$ satisfy the same SDE and we can thus assume in a suitable probability space

$$(3.3) \quad \tilde{Y}_t = \log(\tilde{X}_t/F(t)).$$

In this paper, we call \tilde{X}_t as the *Lognormal Mixture Diffusion* and \tilde{Y}_t as the *Normal Mixture Diffusion*; we call $\tilde{L}_t(\cdot)$ as the *Lognormal Mixture Distribution* and $\tilde{N}_t(\cdot)$ as the *Normal Mixture Distribution*. The equality (3.3) implies that following two modeling approach yield an identical diffusion process: A) the moneyness of the lognormal mixture diffusion whose marginal density function is modeled to match exactly the risk-neutral distribution of asset price; B) the normal mixture diffusion whose marginal density function is modeled to match exactly the risk-neutral distribution of the moneyness of asset price. In this sense, we can model the risk-neutral distribution with lognormal mixture diffusion and normal mixture diffusion interchangeably. This works for mixture of geometric diffusions when the mixing function is time-independent, regardless whether the drift and volatility terms are random or deterministic.

Finally, we present some useful properties for the Fourier transform of the normal mixture distribution. Fourier transform of an absolute integral function $f : \mathbb{R} \mapsto \mathbb{R}$ is defined as

$$\mathcal{F}(f)(\eta) := \int_{-\infty}^{\infty} e^{i\eta x} f(x) dx$$

Lemma 3.1 *For mixture of geometric diffusion (3.1), the Fourier transform of the normal mixture distribution is*

$$(3.4) \quad \mathcal{F}(\tilde{N}_t)(\eta) = \int_{\Theta} m(\theta) \exp \left(i\eta u(\theta, t) - (i\eta + \eta^2)v(\theta, t)/2 \right) d\theta.$$

where $u(\cdot)$ and $v(\cdot)$ are defined in (3.2). Particularly, if the lognormal mixture distribution equals the risk-neutral density function at T , then

$$(3.5) \quad \mathcal{F}(\tilde{N}_T)(-i) = \int_{\Theta} m(\theta) e^{u(\theta, T)} d\theta = 1.$$

Note that the equation (3.5) is consistent with the constraint we derived in (2.19) as the result of the linear growth condition.

3.2 Re-Parametrization with Time-Dependent Mixing Function

In the remaining of this paper, we focus on deriving a suitable parametrization for mixture of geometric diffusion, such that its marginal density functions fit exactly to risk-neutral distribution across all maturities. Our modeling approach generally follow these steps: first we choose a generic parametrization for GGBM, next we solve for the mixing function m_T , such that the resulting mixture distribution equals the given risk-neutral distribution at maturity T . After repeating the second step for every known $T \in (0, T_0]$, we obtain a time-dependent mixing function $\{m_t(\cdot), 0 < t \leq T_0\}$.

Since the interchangeability between normal mixture diffusion and lognormal mixture diffusion requires that the mixing function is time-independent, we will re-parametrize the mixture diffusion and convert the mixing function into a time-independent one. This re-parametrization is achievable with the next proposition, where we can specify an arbitrary time-independent mixing function, then re-parametrize the mixture diffusion such that the re-parametrized mixture diffusion still have the same marginal distribution. This result is applicable to the mixture of geometric diffusion when Θ is one-dimensional. Note that the target mixing function can be set as either a time-dependent function or a time-independent function.

Theorem 3.2 *For the mixture of geometric diffusion (3.1), we assume*

- C1. Θ is a Lebesgue measurable set in \mathbb{R} , $m_t(\theta)$, $\hat{m}_t(\theta) : \Theta \times [0, T_0] \mapsto \mathbb{R}$ are continuous in (t, θ) .*

C2. There exists a continuous function $h : \Theta \rightarrow [0, \infty)$ such that $m_t(\theta) + h(\theta)$ and $\hat{m}_t(\theta) + h(\theta)$ are strictly positive.

C3. Denote the cumulative distribution function of $m_t(\theta) + h(\theta)$ as $M_t(\theta)$ and the cumulative distribution function of $\hat{m}_t(\theta) + h(\theta)$ as $\hat{M}_t(\theta)$. We assume $M_t^{-1}(\hat{M}_t(\theta))$ is differentiable in θ ; $\hat{\mu}(\theta, t)$ and $\hat{\nu}(\theta, t)$ are well-defined, where

$$\begin{aligned}\hat{\mu}(\theta, t) &:= \frac{\partial}{\partial t} \int_0^t \mu(M_t^{-1}(\hat{M}_t(\theta)), s) ds \\ \hat{\nu}(\theta, t) &:= \frac{\partial}{\partial t} \int_0^t \nu(M_t^{-1}(\hat{M}_t(\theta)), s) ds\end{aligned}$$

Parametrize the diffusion $\check{X}_t(\theta, \psi)$ with $(\theta, \psi) \in \Theta \times \{0, 1\}$ as

$$(3.6) \quad \frac{d\check{X}_t(\theta, \psi)}{\check{X}_t(\theta, \psi)} = \begin{cases} \mu(\theta, t) dt + \sqrt{\nu(\theta, t)} dW_t, & \psi = 0; \\ \hat{\mu}(\theta, t) dt + \sqrt{\hat{\nu}(\theta, t)} dW_t, & \psi = 1. \end{cases}$$

Then the mixture diffusion $\check{X}_t(\theta, \psi)$ with the mixing function

$$(3.7) \quad \check{m}_t(\theta, \psi) = \begin{cases} -h(\theta), & \psi = 0; \\ \hat{m}_t(\theta) + h(\theta), & \psi = 1. \end{cases}$$

has the same mixture distribution as that of $X_t(\theta)$ with the mixing function $m_t(\theta)$, for every $t \in [0, T_0]$.

In Section 4, we will derive the mixing function using Fourier transform and this can result in negative values for $m_t(\theta)$. The control function $h(\theta)$ helps to ensure that there is one-to-one map between $M_t(\theta)$ and $\hat{M}_t(\theta)$ when this occurs.

3.3 Re-Parametrized Mixture Diffusion with Random Volatility

Modeling mixture diffusion with random volatility requires that mixing function to be non-negative. In Theorem 3.2, when both $m_t(\cdot)$, $\hat{m}_t(\cdot)$ are non-negative, we can let $h(\theta) = 0$. Then the part of mixture diffusion (3.6) and (3.7) with $\psi = 0$ vanishes. This results in a much simpler version of Theorem 3.2. Note that the target set $\hat{\Theta}$ in the following corollary can be an arbitrary measurable set in \mathbb{R} .

Corollary 3.3 *We assume the following conditions for the mixture of geometric diffusion (3.1):*

- C1. Θ and $\hat{\Theta}$ are Lebesgue measurable sets in \mathbb{R} , $m_t(\theta) : \Theta \times [0, T_0] \mapsto [0, \infty)$ and $\hat{m}_t(\theta) : \hat{\Theta} \times [0, T_0] \mapsto [0, \infty)$ are continuous in (θ, t) .*
- C2. Denote the cumulative distribution function of $m_t(\theta)$ as $M_t(\theta)$ and the cumulative distribution function of $\hat{m}_t(\theta)$ as $\hat{M}_t(\theta)$. We assume $M_t(\cdot)$ and $\hat{M}_t(\cdot)$ are strictly increasing functions.*
- C3. We assume $M_t^{-1}(\hat{M}_t(\theta))$ is differentiable in θ ; and $\hat{\mu}(\theta, t), \hat{\nu}(\theta, t) : \hat{\Theta} \times [0, T_0] \mapsto \mathbb{R}$ are well-defined, where*

$$(3.8) \quad \begin{aligned} \hat{\mu}(\theta, t) &:= \frac{\partial}{\partial t} \int_0^t \mu(M_t^{-1}(\hat{M}_t(\theta)), s) ds \\ \hat{\nu}(\theta, t) &:= \frac{\partial}{\partial t} \int_0^t \nu(M_t^{-1}(\hat{M}_t(\theta)), s) ds \end{aligned}$$

Parametrize diffusion $\hat{X}_t(\theta)$ for $\theta \in \hat{\Theta}$ as

$$(3.9) \quad \frac{d\hat{X}_t(\theta)}{\hat{X}_t(\theta)} = \hat{\mu}(\theta, t) dt + \sqrt{\hat{\nu}(\theta, t)} dW_t.$$

Then the mixture diffusion $\hat{X}_t(\theta)$ with the mixing function $\hat{m}_t(\cdot)$ has the same mixture distribution as that of $X_t(\theta)$ with the mixing function $m_t(\theta)$, for every $t \in (0, T_0]$.

Next we apply Corollary 3.3 to re-parametrize one particularly useful mixture diffusion

$$(3.10) \quad \frac{dX_t(\theta)}{X_t(\theta)} = r(t) dt + \sqrt{\theta} dW_t$$

with the time-dependent mixing function $m_t(\cdot) : (0, \infty) \mapsto [0, \infty)$. Assume the target mixing function is time-independent and denoted it as $m_*(\cdot) : (0, \infty) \mapsto [0, \infty)$. Denote $\hat{\nu}(\theta, t)$ as the re-parametrized volatility function from Corollary 3.3 for the pair of mixing functions (m_t, m_*) . We write the resulting mixture diffusion with random volatility derived in Theorem 2.4 as

$$(3.11) \quad \frac{d\tilde{X}_t}{\tilde{X}_t} = r(t) dt + \sqrt{\hat{\nu}(V, t)} dW_t,$$

where V is a random variable with the probability density function $m_*(\cdot)$.

Note that the re-parametrized volatility term $\hat{\nu}(\theta, t)$ in (3.11) can be alternatively expressed as

$$(3.12) \quad M_t(\bar{\nu}(\theta, t)) = M_*(\theta)$$

where $\bar{\nu}(\theta, t) := \frac{1}{t} \int_0^t \hat{\nu}(\theta, s) ds$ is the annualized volatility; M_t and M_* are the cumulative distribution function of m_t and m_* , respectively. The re-parametrization (3.12) essentially maps the annualized volatility via the monotonic function M_t to the cumulative distribution function we have picked, i.e., \hat{M}_* .

Some contingent claims such as European options and variance swaps depend only on the annualized volatility. The following corollary shows it suffices to calculate these claims under the original mixing function $m_t(\cdot)$.

Corollary 3.4 *Assume the payoff function of the contingent claim $C(\tilde{X})$ for the mixture diffusion (3.11) can be also expressed as $f(\bar{\nu}(V, T))$, where $f(\cdot) : \mathbb{R} \mapsto \mathbb{R}$ is a measurable function. Then the expected value of the contingent claim is*

$$\mathbb{E}C(\tilde{X}) = \int_0^\infty f(\theta) m_T(\theta) d\theta.$$

Finally, we show there is simple way to verify the conditions of Theorem 2.4 for the mixture diffusion (3.11).

Corollary 3.5 *Assume there exists a positive constant K such that*

$$\hat{\nu}(\theta, t) \leq K \int_0^{T_0} \hat{\nu}(\theta, s) ds.$$

Then the sufficient condition such that the mixture diffusion (3.11) admits a unique strong solution is

$$\int_{\Theta} e^{2C_0 T_0 K \theta} m_{T_0}(\theta) d\theta < \infty.$$

4 Explicit Mixture Diffusion Models

In this section, we derive the explicit formula of the mixing function based on various parameterization approaches such that the normal mixture distribution equals the risk-neutral distribution of the moneyness of asset price. We denote the risk-neutral distribution at time t for the moneyness of asset price as $E_t(\cdot)$ and assume $E_t(\cdot)$ is continuous.

4.1 Risk-Neutral Mixture Diffusion with Random Volatility

We consider the following parametrization of mixture of geometric diffusion with random volatility

$$(4.1) \quad \frac{dX_t(\theta)}{X_t(\theta)} = r(t) dt + \sqrt{\theta} dW_t$$

where $\theta \in \mathbb{R}^+$. The following Lemma shows we can use Laplace inversion to solve the mixing function from the risk-neutral distribution. The unilateral *Laplace transform* for a suitable function $f : [0, \infty) \mapsto \mathbb{R}$ is defined as

$$\mathcal{L}(f)(t) = \int_0^\infty e^{-tx} f(x) dx$$

The *Inverse Laplace Transform* of a function $F(s)$ is the piece-wise continuous and exponentially-restricted real function $f : \mathbb{R} \mapsto \mathbb{R}$ such that $\mathcal{L}(f)(t) = F(t)$. We denote the Laplace inversion of $F(\cdot)$ as $\mathcal{L}^{-1}(F)(t)$.

Lemma 4.1 *The necessary and sufficient condition such that the mixture diffusion (4.1) with mixing function $m_T : (0, \infty) \mapsto [0, \infty)$ has the normal mixture distribution $E_T(\cdot)$ is that, $G_T(\cdot)$ is completely monotonic, where*

$$G_T(\eta) := \mathcal{F}(E_T) \left(\sqrt{\frac{2\eta}{T}} - \frac{1}{4} - \frac{i}{2} \right)$$

Then the Laplace inversion of G_T exists and $m_T(x) = \mathcal{L}^{-1}(G_T)(x)$.

It is natural to extend (4.1) to a multi-factor mixture diffusion

$$(4.2) \quad \frac{dX_t(\vec{\theta})}{X_t(\vec{\theta})} = r(t) dt + \sum_{k=1}^n \sqrt{\nu^{(k)}(\theta^{(k)}, t)} dW_t^{(k)}$$

where $(W_t^{(1)}, \dots, W_t^{(n)})$ is a vector of independent standard Brownian motions and $\vec{\theta} = (\theta^{(1)}, \dots, \theta^{(n)}) \in (0, \infty)^n$. Both (4.1) and (4.2) can be used to solve the risk-neutral distribution $E_t(\cdot)$. The following proposition establishes the relation between the mixing functions of these solutions.

Proposition 4.2 *For every $k = 1, \dots, n$, denote $M_t^{(k)}$ as the cumulative distribution function of the density function $m_t^{(k)} : (0, \infty) \mapsto [0, \infty)$ and $\bar{\nu}^{(k)}(x, t) := \frac{1}{t} \int_0^t \nu^{(k)}(x, s) ds$. We assume the inverse of $\bar{\nu}^{(k)}(\cdot, t)$ exists for every $t \in (0, T_0]$ and denote it as $\check{\nu}^{(k)}(\cdot, t)$.*

Assume the mixture distribution of (4.2) with the mixing function $\prod_{k=1}^n m_t^{(k)}(\theta^{(k)})$ equals that of (4.1) with the mixing function $m_t : (0, \infty) \mapsto [0, \infty)$. Then we have

$$m_t = \check{m}_t^{(1)} * \check{m}_t^{(2)} * \dots * \check{m}_t^{(n)}$$

where $\check{m}_t^{(k)}(x) = \frac{\partial}{\partial x} M_t^{(k)}(\check{\nu}^{(k)}(x, t))$.

4.2 Risk-Neutral Mixture Diffusion with Stochastic Volatility

In this section we will parametrize the drift term and convert it into something related to stochastic volatility. Consider the parametrization

$$(4.3) \quad \frac{dX_t(\theta, \psi)}{X_t(\theta, \psi)} = (r(t) + \theta) dt + \sqrt{\nu(\psi, t)} dW_t$$

where the drift and volatility parameters have independent mixing distribution, i.e., $m(\theta)b(\psi)$; $b(\cdot) : (0, \infty) \mapsto [0, \infty)$ is a given density function and $\theta \in \mathbb{R}$. The following lemma gives the necessary and sufficient condition to solve the mixing function $m(\cdot)$.

Lemma 4.3 *The necessary and sufficient condition such that the mixture diffusion (4.3) with the mixing function $m(\theta)b(\psi)$ has the normal mixture distribution $E_T(\cdot)$ is that,*

$G_T(\cdot)$ is positive definite, where

$$G_T(\eta) = \frac{\mathcal{F}(E_T)(\eta)}{\int_0^\infty b(\psi) \exp\left(-\frac{1}{2}(i\eta + \eta^2) \int_0^T \nu(\psi, s) ds\right) d\psi}.$$

Then we have $m(x) = \frac{1}{T} \mathcal{F}^{-1}(G_T)\left(\frac{x}{T}\right)$.

Next, we show how to convert the drift term into a diffusion process. Define

$$F_t(x) := \frac{t}{x} m_t(t \log(x))$$

Heuristically, if we regard $F_t(\cdot)$ as a *risk-neutral* distribution function, then deriving the risk-neutral diffusion for $E_t(\cdot)$ is equivalent to solving the *risk-neutral* diffusion for $F_t(\cdot)$.

Here we assume the risk-neutral diffusion $\{\tilde{U}_t, 0 \leq t \leq T_0\}$:

$$(4.4) \quad \begin{aligned} \frac{d\tilde{U}_t}{\tilde{U}_t} &= \sqrt{\beta(\tilde{U}_t, t)} dB_t \\ \tilde{U}_0 &= 1 \end{aligned}$$

is such a solution and it has the marginal density function $F_t(\cdot)$ for every $0 < t \leq T_0$.

Finally, the following corollary shows that the desired risk-neutral diffusion consists of three parts: a deterministic drift term $r(t)$, a stochastic volatility term $\beta(\tilde{U}_t, t)$ and a random volatility term $\nu(V, t)$.

Corollary 4.4 *Assume the unique strong solution exists for the diffusion*

$$(4.5) \quad \frac{d\tilde{X}_t}{\tilde{X}_t} = r(t) dt + \sqrt{\beta(\tilde{U}_t, t)} dB_t + \sqrt{\nu(V, t)} dW_t$$

where \tilde{U}_t is the diffusion process defined in (4.4) and V is a random variable with density function $b(\cdot) : (0, \infty) \mapsto [0, \infty)$. V , B_t and W_t are independent to each other.

Then \tilde{X}_t has the marginal density function $E_t(\cdot)$ for every $0 < t \leq T_0$.

Approaches in Section 4.1 and 4.3 can lead to explicit solutions for \tilde{U}_t . In the case that \tilde{U}_t is a mixture diffusion with random volatility, then (4.5) is reduced to a special case of two-factor random volatility model. In the case that \tilde{U}_t is a mixture diffusion with deterministic volatility, then the volatility term of the diffusion process is stochastic.

4.3 Mixture Diffusion with Deterministic Drift and Volatility

In this section, we present various one-dimensional parametrizations of geometric diffusion (3.1) that leads to straightforward solution for a risk-neutral distribution. Because these parametrizations can lead to negative values in mixing function, they are applicable only for mixture diffusion with deterministic drift and volatility.

We start with the parametrization used in Lemma 4.1, i.e.,

$$(4.6) \quad \frac{dX_t(\theta)}{X_t(\theta)} = r(t) dt + \sqrt{\theta} dW_t.$$

With some modifications, we can solve the mixing function under much less restrictive conditions. Our new parametrization is defined over $(\theta, \psi) \in \mathbb{R}^+ \times \{+, -\}$: when $\psi = +$, we define it as (4.6); when $\psi = -$, we define it as the inversion $F^2(t)/X_t(\theta)$, i.e.,

$$(4.7) \quad \begin{aligned} \frac{dX_t(\theta, \psi)}{X_t(\theta, \psi)} &= \begin{cases} r(t) dt + \sqrt{\theta} dW_t, & \psi = + \\ (r(t) + \theta) dt - \sqrt{\theta} dW_t, & \psi = - \end{cases} \\ X_0(\theta, \psi) &= x_0 \end{aligned}$$

The following proposition solves the mixing function for (4.7).

Proposition 4.5 *Define $H_\psi(x)$ for $x > 0$ and $\psi \in \{+, -\}$ as*

$$H_\psi(x^2) = \frac{e^{x/2} E_T(-\psi x) - e^{-x/2} E_T(\psi x)}{e^x - e^{-x}}$$

We assume the Laplace inversions for both $H_+(\cdot)$ and $H_-(\cdot)$ exist. Then the mixture diffusion (4.7) with the mixing function

$$m(\theta, \psi) = \sqrt{\frac{\pi}{2\theta^3 T}} \exp\left(\frac{\theta T}{8}\right) \mathcal{L}^{-1}(H_\psi)\left(\frac{1}{2\theta T}\right)$$

has the normal mixture distribution $E_T(\cdot)$ at time T .

Next we consider the parametrization of the drift term with the constant volatility

$$(4.8) \quad \frac{dX_t(\theta)}{X_t(\theta)} = (r(t) + \theta) dt + \sqrt{\nu_0} dW_t,$$

where $\theta \in \mathbb{R}$. The following proposition shows we can solve the mixing function using Fourier transform.

Proposition 4.6 *We assume*

$$g(\eta) := \exp\left((i\eta + \eta^2)T\nu_0/2\right)\mathcal{F}(E_T)(\eta) \in \mathcal{L}^1(\mathbb{R}).$$

Then the mixture diffusion (4.8) with the mixing function

$$m_T(\theta) = \frac{1}{2\pi T} \Re\left(\mathcal{F}(g)\left(-\frac{\theta}{T}\right)\right)$$

has the normal mixture distribution $E_T(\cdot)$ at time T .

In fact the parametrization (4.8) is not only the simplest way to parametrize the drift term, but also contains all possible parametrizations of geometric diffusions when their total variance have a none-zero lower bound.

Corollary 4.7 *Assume there exists $V_0 > 0$ such that for every $\theta \in \Theta$, the total variance of the mixture of geometric diffusion (3.1) satisfies*

$$(4.9) \quad \int_0^T \nu(\theta, t) dt > V_0.$$

Then there exists a mixture diffusion as parametrized in (4.8), which has the same normal mixture distribution as that of the mixture diffusion (3.1) at time T .

Many stochastic volatility models can have a none-zero probability such that the total variance stays below any arbitrary small positive value [Hagan *et al* (2002), Heston (1993)]. This will violate the assumption in (4.9). We can overcome this problem by assigning a density function to the volatility term. Define our new parametrization as

$$(4.10) \quad \begin{aligned} \frac{dX_t(\theta, \psi)}{X_t(\theta, \psi)} &= (r(t) + \theta + \psi/2) dt + \sqrt{\psi} dW_t. \\ X_0(\theta, \psi) &= x_0 \end{aligned}$$

where $(\theta, \psi) \in \mathbb{R} \times \mathbb{R}^+$. The following proposition explicitly solves the mixing function for the parametrization (4.10).

Proposition 4.8 *We assume $b(\cdot) : (0, \infty) \mapsto [0, \infty)$ is a known density function. Assume*

$$(4.11) \quad g(\eta) := \frac{\mathcal{F}(E_T)(\eta)}{\mathcal{L}(b)(\eta^2 T/2)} \in \mathcal{L}^1(\mathbb{R}).$$

Then the mixture diffusion (4.10) with the mixing function

$$(4.12) \quad \frac{b(\psi)}{2\pi T} \Re \left(\mathcal{F}(g) \left(-\frac{\theta}{T} \right) \right)$$

has the normal mixture distribution $E_T(\cdot)$ at time T .

We can choose a suitable function $b(\cdot)$ so that the conditions (4.11) can be generally satisfied for most of applications. For example, if we set $b(\cdot)$ as the Gamma density function, i.e., $b(x) = \frac{1}{\Gamma(\alpha)} \beta^\alpha x^{\alpha-1} e^{-x\beta}$, then (4.11) becomes

$$(1 + \eta^2 T/2)^\alpha \mathcal{F}(E_T)(\eta) \in \mathcal{L}^1(\mathbb{R}).$$

5 Term Structure Model

5.1 Solution of Mixture Diffusion

We summarize our approach in modeling mixture diffusion from the previous three sections: First, we calibrate a time-dependent mixing function such that the mixture distributions equals the risk-neutral distributions across all maturities. Section 4 details several approaches to achieve this. Next, we choose a time-independent mixing function and apply techniques in Section 3 to re-parametrize the drift and volatility function along the maturity. The mixture distribution of the re-parametrized diffusion will remain the same as that of the original mixture diffusion. Finally, we apply the results from Section 2 to obtain a valid diffusion process with unique strong solution. Below we loosely define the resulting mixture diffusion as the *Solution of Mixture Diffusion*.

Definition 5.1 *We say $(\hat{m}, \hat{\nu}(\hat{V}, t), \hat{\mu}(\hat{V}, t))$ is the solution of the mixture diffusion (3.1) for risk-neutral distribution $D_t(\cdot)$ if it satisfies the following conditions:*

C1. For every $0 < T \leq T_0$, there exists a mixing function $m_T(\cdot)$ such that the mixture distribution of the mixture of geometric diffusion (3.1) with mixing function $m_T(\cdot)$ equals the risk-neutral distribution $D_T(\cdot)$.

C2. $\hat{m}(\cdot)$ is a density function. Corollary 3.3 holds true for the re-parametrization between the pair (m_t, \hat{m}) . Denote $\hat{\mu}(\cdot)$ and $\hat{\nu}(\cdot)$ as the re-parametrized drift and volatility terms defined in (3.8), respectively.

C3. Let \hat{V} be the random variable with the density function $\hat{m}(\cdot)$. The mixture diffusion with random drift and volatility

$$(5.1) \quad \frac{d\hat{X}_t}{\hat{X}_t} = \hat{\mu}(\hat{V}, t) dt + \sqrt{\hat{\nu}(\hat{V}, t)} dW_t.$$

admits a unique strong solution and has the mixture distribution $D_t(\cdot)$ for every $t \in (0, T_0]$.

In practice, we only have sufficient information to reasonably estimate of the risk-neutral distributions at few maturities where the European options are actively traded. For the rest of maturities, we need find a way to fill the data void. There are many methods to achieve it and they mostly use the interpolated risk-neutral function, mixture function, volatility function, or a combination of them. In this sense, when the risk-neutral distribution is available only at a finite number of maturities, we refer the solution of the mixture diffusion to the case that the risk-neutral functions are populated with a pre-determined interpolation method. The solution is then uniquely determined by the pre-determined interpolation method.

5.2 Modeling Instantaneous Forward Rate

In this section, we will derive the mixture diffusion model of instantaneous forward rate such that it can price Cap/Floor options consistently with the market. Denote $B(t, T)$ as the price of zero-coupon bond with maturity T at the forward time t . The relation

between the instantaneous forward rate $f(t, T)$ and zero-coupon bond is defined via the equation,

$$B(t, T) = e^{-\int_t^T f(t, u) du}.$$

The generic forward LIBOR with expiry T_1 and maturity T_2 at forward time t is defined as

$$L(t, T_1, T_2) := \frac{1}{T_2 - T_1} \left(\frac{B(t, T_1)}{B(t, T_2)} - 1 \right)$$

where $0 \leq t \leq T_1 < T_2 \leq T_0$.

The proposed mixture diffusion model for the instantaneous forward rate is calibrated to the risk-neutral distributions of a series of LIBOR with consecutive expiry and maturity pairs, i.e.,

$$L(t, T_{(0)}, T_{(1)}), L(t, T_{(1)}, T_{(2)}), \dots, L(t, T_{(n-1)}, T_{(n)})$$

where $0 < T_{(0)} < T_{(1)} < \dots < T_{(n)} \leq T_0$. We denote $D_t^{(k)}(\cdot)$ as the risk-neutral distribution of $L(t, T_{(k-1)}, T_{(k)})$ and $E_t^{(k)}(\cdot)$ as that of the *moneyness* of LIBOR, i.e.,

$$(5.2) \quad -\log((T_{(k)} - T_{(k-1)})L(t, T_{(k-1)}, T_{(k)}) + 1).$$

Then it is clear $E_t^{(k)}(\cdot)$ and $D_t^{(k)}(\cdot)$ satisfies

$$E_t^{(k)}(x) = \frac{e^{-x}}{T_{(k)} - T_{(k-1)}} D_t^{(k)} \left(\frac{e^{-x} - 1}{T_{(k)} - T_{(k-1)}} \right).$$

In practice, the underlying of the commonly traded LIBOR options is the spot LIBOR with t equals the expiry, i.e., $L(T_{(k-1)}, T_{(k-1)}, T_{(k)})$. As discussed previously, we can populate the risk-neutral distribution for the rest of forward time with some suitable interpolation method.

We assume that the scalar function $\hat{\phi}^{(k)}(t, T) : [0, T_0] \times [0, T_0] \mapsto \mathbb{R}$ has continuous derivation in T . It satisfies $\hat{\phi}^{(k)}(t, T) = 0$ for $T \leq t$ and

$$(5.3) \quad \hat{\phi}^{(k)}(t, T_{(m)}) = \begin{cases} 1, & m \geq k \text{ and } t \leq T_{(k-1)} \\ 0, & m \leq k - 1 \text{ and } t \leq T_m \end{cases}$$

for every $1 \leq m \leq n$. One example of such function is

$$(5.4) \quad \hat{\phi}^{(k)}(t, T) = \frac{f(T, T_{(k-1)}, T_{(k)})f(t, T, -1)}{f(\min(t, T_{(k-1)}), T'_{(k)} + f(T, T'_{(k)}, T_{(k)})(T - T'_{(k)}), -1)}$$

where $T'_{(k)} \in (T_{(k-1)}, T_{(k)})$ is a fixed number and

$$(5.5) \quad f(x, a, b) = \begin{cases} 0, & (x - a)(b - a) \leq 0 \\ \sin^2\left(\frac{\pi}{2} \frac{x - a}{b - a}\right), & \min(a, b) < x < \max(a, b) \\ 1, & (x - b)(b - a) \geq 0 \end{cases}$$

It is clear that the denominator in (5.4) is always positive. Therefore, $\hat{\phi}^{(k)}(t, T)$ has continuous derivation in T . Since $f(t, T, -1) = 0$ for $T \leq t$, we have $\hat{\phi}^{(k)}(t, T) = 0$ as well. When $t \leq T_{(k-1)}$ and $T \geq T_{(k)}$, the denominator cancels out $f(t, T, -1)$. Consequently, we have $\hat{\phi}^{(k)}(t, T) = 1$.

Below is our main result for the term structure model of the instantaneous forward rate.

Theorem 5.2 *Let $W_t = (W_t^{(1)}, \dots, W_t^{(n)})$ be a vector of independent standard Brownian motions. We assume for every $k = 1, \dots, n$, the trio $(\hat{m}^{(k)}, \hat{\nu}^{(k)}(\hat{V}^{(k)}, t), 0)$ solves the mixture diffusion*

$$(5.6) \quad \begin{aligned} \frac{dX^{(k)}(t; \theta)}{X^{(k)}(t; \theta)} &= \sqrt{\nu^{(k)}(\theta, t)} dW_t^{(k)} \\ X^{(k)}(0; \theta) &= B(T_{(k)})/B(T_{(k-1)}) \end{aligned}$$

for the risk-neutral distribution $E_t^{(k)}(\cdot)$ of the moneyness of the forward LIBOR. Define the SDE for the instantaneous forward rate as

$$(5.7) \quad \begin{aligned} d\hat{f}(t, T) &= \frac{1}{2} \sum_{k=1}^n \frac{\partial(\hat{\phi}^{(k)}(t, T))^2}{\partial T} \hat{\nu}^{(k)}(\hat{V}^{(k)}, t) dt + \sum_{k=1}^n \frac{\partial \hat{\phi}^{(k)}(t, T)}{\partial T} \sqrt{\hat{\nu}^{(k)}(\hat{V}^{(k)}, t)} dW_t^{(k)} \\ \hat{f}(0, T) &= f_0(T) \end{aligned}$$

where $f_0(T)$ is the initial instantaneous forward rate and $\hat{\phi}^{(k)}(t, T)$ is defined in (5.3).

Then the zero-coupon bond process $\hat{B}(t, T) = e^{-\int_t^T \hat{f}(t, u) du}$ is risk-neutral and the forward LIBOR process

$$\hat{L}(t, T_{(k-1)}, T_{(k)}) := \frac{e^{\int_{T_{(k-1)}}^{T_{(k)}} \hat{f}(t, u) du} - 1}{T_{(k)} - T_{(k-1)}}$$

has the marginal density function $D_t^{(k)}(\cdot)$ for every $k = 1, \dots, n$ and $0 < t \leq T_{(k-1)}$.

As an application of Theorem 5.2, we consider pricing the *Option on Interest Rate Swaps* (Swaption) with our term structure model of instantaneous forward rate. Particularly, we consider a *Payer Forward-start Interest-Rate Swap* (PFS) which receives the floating rate $L(t, T_{(k-1)}, T_{(k)})$ for the period $(T_{(k-1)}, T_{(k)}]$ and pays the constant fixed rate K , where $k = \alpha + 1, \dots, \beta$. Then the payoff function of discounted PFS, denoted as $\hat{P}(t, K)$ is,

$$\begin{aligned} \hat{P}(t, K) &= B^{-1}(t) \sum_{k=\alpha+1}^{\beta} (T_{(k)} - T_{(k-1)}) (\hat{L}(t, T_{(k-1)}, T_{(k)}) - K) \hat{B}(t, T_{(k)}) \\ &= B^{-1}(t) \hat{B}(t, T_{(\alpha)}) - B^{-1}(t) \hat{B}(t, T_{(\beta)}) - K \sum_{k=\alpha+1}^{\beta} (T_{(k)} - T_{(k-1)}) B^{-1}(t) \hat{B}(t, T_{(k)}) \\ (5.8) \quad &:= \sum_{k=\alpha}^{\beta} c_{(k)}(K) B^{-1}(t) \hat{B}(t, T_{(k)}) \end{aligned}$$

The following proposition derives the price of the swaption with the payoff function (5.8).

Proposition 5.3 Assume $\hat{B}(t, T)$ is the zero-coupon bond process in Theorem 5.2. When $0 < t \leq T_{(\alpha)}$, the value of swaption based on the discounted PFS (5.8) with the strike rate K is

$$\mathbb{E}[\hat{P}(t, K)]^+ = \mathbb{E} \left[1 + \sum_{m=\alpha+1}^{\beta} c_{(m)}(K) \exp \left(\sum_{k=\alpha+1}^m -\frac{t}{2} \bar{V}_t^{(k)} + \sqrt{t \bar{V}_t^{(k)}} N^{(k)} \right) \right]^+$$

where $\bar{V}_t^{(k)} = \frac{1}{t} \int_0^t \hat{v}^{(k)}(\hat{V}^{(k)}, s) ds$. $N = (N^{(1)}, \dots, N^{(n)})$ is a vector of i.i.d. standard normal random variables and $\hat{V} = (\hat{V}^{(1)}, \dots, \hat{V}^{(n)})$ is defined in Theorem 5.2. N and \hat{V} are independent to each other.

Theorem 5.2 depends only on the marginal density function of the random vector \hat{V} . However, the swaption price of Proposition 5.3 depends on the joint distribution of \hat{V} . Therefore, the joint distribution of \hat{V} can be calibrated to the contingent claims such as swaptions.

5.3 Modeling Nested LIBORs

In our term structure model of instantaneous forward rate, we assume there is no overlapping among the expiry and maturity pairs of forward LIBORs. However, there usually exist multiple maturities for the same expiry. This leads to the scenario that one expiry and maturity pair completely nests inside another pair. To adapt our term structure to this scenario, we consider the general setting where two adjacent LIBORs nest inside another LIBOR. To be more specific, we consider the case that $L(t, T_{(k-1)}, T_{(k-.5)})$ and $L(t, T_{(k-.5)}, T_{(k)})$ nest inside $L(t, T_{(k-1)}, T_{(k)})$, where $T_{(k-1)} < T_{(k-.5)} < T_{(k)}$ and $1 \leq k \leq n$. We will create additional models for $L(t, T_{(k-1)}, T_{(k-.5)})$ and $L(t, T_{(k-.5)}, T_{(k)})$ such that they are consistent with the existing term structure model in Theorem 5.2. When the number of the nested LIBORs is more than two, we can first bundle them into two non-overlapping groups, then model each group. We will repeat this step until there is no more than two nested LIBORs inside every expiry and maturity pair.

Since Cap/Floor options are based on spot LIBOR, we only need to model $L(t, T_{(k-1)}, T_{(k-.5)})$ and $L(t, T_{(k-.5)}, T_{(k)})$ at their expiries, i.e., $t = T_{(k-1)}$ and $T_{(k-.5)}$. We denote $D_A^{(k)}(\cdot)$ and $D_B^{(k)}(\cdot)$ as the risk-neutral function of $L(T_{(k-1)}, T_{(k-1)}, T_{(k-.5)})$ and $L(T_{(k-.5)}, T_{(k-.5)}, T_{(k)})$, respectively. The risk-neutral function of the *moneyness* of the forward LIBORs are denoted as $E_A^{(k)}(\cdot)$ and $E_B^{(k)}(\cdot)$, respectively.

Consider the following two integrals of the instantaneous forward processes in Theorem 5.2,

$$\int_{T_{(k-1)}}^{T_{(k-.5)}} \hat{f}(t, u) du \quad \text{and} \quad \int_{T_{(k-.5)}}^{T_{(k)}} \hat{f}(t, u) du.$$

Denote their marginal density functions as $\hat{F}_A^{(k)}(\cdot)$ and $\hat{F}_B^{(k)}(\cdot)$, respectively. We assume

there exist some density functions $\tilde{F}_A^{(k)}(\cdot)$ and $\tilde{F}_B^{(k)}(\cdot)$ such that

$$(5.9) \quad \hat{F}_A^{(k)} * \tilde{F}_A^{(k)} = E_A^{(k)}, \quad \hat{F}_B^{(k)} * \tilde{F}_B^{(k)} = E_B^{(k)}.$$

In the following theorem, we will assume that the function $\tilde{\phi}^{(k)}(t, T) : [0, T_0] \times [0, T_0] \mapsto \mathbb{R}$ has continuous derivation in T . It satisfies $\tilde{\phi}^{(k)}(t, T) = 0$ for $T \leq t$ and

$$(5.10) \quad \tilde{\phi}^{(k)}(t, T_{(m)}) = \begin{cases} 1, & m = k - .5 \text{ and } t \leq T_{(k-1)} \\ 0, & m \neq k - .5 \text{ and } t \leq T_m \end{cases}$$

for every $1 \leq m \leq n$. One example of such function is

$$\tilde{\phi}^{(k)}(t, T) = \frac{f(T, T_{(k-1)}, T_{(k-.5)})f(T, T_{(k)}, T_{(k-.5)})f(t, T, -1)}{f(\min(t, T_{(k-1)}), T'_{(k)} + f(T, T'_{(k)}, T_{(k-.5)})(T - T'_{(k)}), -1)}$$

where $T'_{(k)} \in (T_{(k-1)}, T_{(k-.5)})$ is a fixed number and $f(\cdot)$ is defined in (5.5).

The following result shows the nested LIBOR can be modeled by an additional component of instantaneous forward rate.

Proposition 5.4 *Let $\tilde{W}_t = (\tilde{W}_t^{(1)}, \dots, \tilde{W}_t^{(n)})$ be a vector of independent standard Brownian motions. We assume for every $k = 1, \dots, n$, the trio $(\tilde{m}^{(k)}, \tilde{\nu}^{(k)}(\tilde{V}^{(k)}, t), 0)$ solves the mixture diffusion*

$$(5.11) \quad \begin{aligned} \frac{d\tilde{X}^{(k)}(t; \theta)}{\tilde{X}^{(k)}(t; \theta)} &= \sqrt{\nu^{(k)}(\theta, t)} d\tilde{W}_t^{(k)} \\ \tilde{X}^{(k)}(0; \theta) &= 1 \end{aligned}$$

for the risk-neutral distribution $\tilde{F}_A^{(k)}(x)$ at $t = T_{(k-1)}$ and $\tilde{F}_B^{(k)}(-x)$ at $t = T_{(k-.5)}$, where $\tilde{F}_A^{(k)}(\cdot)$ and $\tilde{F}_B^{(k)}(\cdot)$ are defined in (5.9).

Let $\hat{f}(t, T)$ be the instantaneous forward rate from in Theorem 5.2. Define

$$\begin{aligned} d\tilde{f}(t, T) &= \frac{1}{2} \sum_{k=1}^n \frac{\partial(\tilde{\phi}^{(k)}(t, T))^2}{\partial T} \tilde{\nu}^{(k)}(\tilde{V}^{(k)}, t) dt + \sum_{k=1}^n \frac{\partial \tilde{\phi}^{(k)}(t, T)}{\partial T} \sqrt{\tilde{\nu}^{(k)}(\tilde{V}^{(k)}, t)} d\tilde{W}_t^{(k)} \\ \tilde{f}(0, T) &= 1 \end{aligned}$$

where $\tilde{\phi}^{(k)}(t, T)$ is defined in (5.10). We assume $\tilde{W}_t, W_t, \tilde{V}, V$ are independent to each other.

Then the zero-coupon bond process $\tilde{B}(t, T) = e^{-\int_t^T \hat{f}(t, u) + \tilde{f}(t, u) du}$ is risk-neutral and the forward LIBOR process

$$\tilde{L}(t, T_{(k-1)}, T_{(k)}) := \frac{e^{\int_{T_{(k-1)}}^{T_{(k)}} \hat{f}(t, u) + \tilde{f}(t, u) du} - 1}{T_{(k)} - T_{(k-1)}}$$

has the marginal density function $D_t^{(k)}(\cdot)$ for every $k = 1, \dots, n$ and $0 < t \leq T_{(k-1)}$.

Furthermore, $\tilde{L}(T_{(k-1)}, T_{(k-1)}, T_{(k-0.5)})$ and $\tilde{L}(T_{(k-0.5)}, T_{(k-0.5)}, T_{(k)})$ have the marginal density function $D_A^{(k)}(\cdot)$ and $D_B^{(k)}(\cdot)$, respectively.

6 Conclusion

This paper shows that with mixture of infinite diffusions, we can derive a diffusion process that not only prices the European options consistently with the market across all maturities and strikes, but also has the explicit formula for any path-dependent options that have the closed-form solution under GGBM. With mixture diffusion, we are able to combine the advantages of many existing asset pricing models into one single framework.

First, mixture diffusion retains the simplicity of the Black-Scholes model and can price options as the weighted average option prices of the underlying GGBM. Most known asset pricing models do not have closed-form solution for path-dependent options and relies on Monte-Carlo simulation or approximation for their valuations. Consequently, their pricing process is neither precise nor efficient. With mixture diffusion, we can efficiently and accurately price many exotic options in consistency with European option market.

Secondly, mixture diffusion has the flexibility of Dupire's local volatility model and we can calibrate it such that its European option prices are consistent with the market across all maturities and strikes. Furthermore, the underlying diffusion processes have explicit parametrization and their parameters can be estimated directly by minimizing some penalty function. Therefore, we can calibrate the mixture diffusion with limited data inputs. In contrast, Dupire's Local Volatility model requires the existence of European option prices across entire volatility surface.

Thirdly, mixture diffusion can capture the randomness of the volatility in the same manner as stochastic volatility models. Though models with stochastic volatility have more realistic asset price movements than those of other asset pricing models, their drawbacks are equally pronounced: stochastic volatility models are not very analytically traceable and some do not even have the closed-form solution for European options. Additionally, stochastic volatility models are parametrized with just very few parameters and they lack the flexibility to fit exactly to European option market. With mixture diffusion, we can retain the advantage of stochastic volatility models, while eliminate these drawbacks by making its European option prices consistent with the market.

Lastly, as an application of mixture diffusion, we model the term structure of forward price under a single risk-neutral measure and produce LIBOR option prices that are consistent with the market for all expiry and maturity pairs. In LIBOR Market Model (or BGM model) [Brace *et al.* (1997), Brigo, Mercurio (2006)], the standard approach is to model every forward LIBOR with an individual Brownian motion and assume these Brownian motions are instantaneous correlated. Compared to LIBOR Market Model, the mixture diffusion can model the underlying instantaneous forward rate directly and produce self-consistent LIBOR processes. Additionally, we can apply the mixture diffusion model to value other contingent claims as we demonstrated with swaptions.

7 Appendix

Proof of Theorem 2.1: We can verify the sufficient conditions in Theorem 2.2 of [Friedman (1975)] for the SDE of $(\tilde{X}_t, \tilde{V}_t)$

$$\begin{aligned} d\tilde{X}_t &= \mu(\tilde{X}_t, t; \tilde{V}_t) dt + \sqrt{\nu(\tilde{X}_t, t; \tilde{V}_t)} dW_t \\ d\tilde{V}_t &= 0 \\ \tilde{X}_0 &= x_0 \\ \tilde{V}_0 &= V \end{aligned}$$

and they are clearly satisfied. Therefore, the unique strong solution exists. Because the

conditional distribution function of \tilde{X}_t given V is $P(x, t; V)$ and V is independent to W_t , the marginal distribution of \tilde{X}_t is $\tilde{P}(x, t)$. ■

Proof of Proposition 2.2: Note that

$$\frac{1}{2} \frac{\partial^2}{\partial x^2} \left((x^2 \beta(x, t; \theta) - \nu(x, t; \theta)) P(x, t; \theta) \right) = \frac{\partial}{\partial x} \left((r(t)x - \mu(x, t; \theta)) P(x, t; \theta) \right)$$

Therefore $P(x, t; \theta)$ satisfies the Fokker-Planck equation for diffusion process

$$\begin{aligned} \frac{dY_t(\theta)}{Y_t(\theta)} &= r(t) dt + \sqrt{\beta(Y_t(\theta), t; \theta)} dW_t \\ Y_0(\theta) &= x_0 \end{aligned}$$

Applying Theorem 2.1 to $Y_t(\theta)$ proves the result. ■

Proof of Corollary 2.3: Because $X_t(\theta)$ is square-integrable, for every $(t, \theta) \in (0, T_0] \times \Theta$, we can find a sequence of positive numbers $\{x_k, k = 1, 2, \dots\}$ such that $\lim_{k \rightarrow \infty} x_k = \infty$ and

$$\lim_{k \rightarrow \infty} x_k^2 P(x_k, t; \theta) = 0$$

Combining the linear growth condition (2.9) with the definition (2.6), we have

$$(7.1) \quad \left| \int_0^x P(x, t; \theta) (r(t)x - \mu(x, t; \theta)) dx \right| \leq K(\theta)(1 + x^2)P(x, t; \theta).$$

Therefore,

$$\begin{aligned} (7.2) \quad & \left| \int_0^\infty P(x, t; \theta) (r(t)x - \mu(x, t; \theta)) dx \right| \\ &= \lim_{k \rightarrow \infty} \left| \int_0^{x_k} P(x, t; \theta) (r(t)x - \mu(x, t; \theta)) dx \right| \\ &\leq \lim_{k \rightarrow \infty} K(\theta)(1 + x_k^2)P(x_k, t; \theta) \\ &= 0. \end{aligned}$$

Since $P(x, t; \theta)$ is the marginal distribution function of $X_t(\theta)$, equivalently, we can write the equation (7.2) as

$$\mathbb{E}(r(t)X_t(\theta) - \mu(X_t(\theta), t; \theta)) = 0.$$

Let $Y_t(\theta) = e^{-\int_0^t r(s) ds} X_t(\theta)$. Then,

$$dY_t(\theta) = e^{-\int_0^t r(s) ds} (-r(t)X_t(\theta) + \mu(X_t(\theta), t; \theta)) dt + e^{-\int_0^t r(s) ds} \sqrt{\nu(X_t(\theta), t; \theta)} dW_t$$

Taking the expectation on both side of the equation, we have

$$d\mathbb{E}Y_t(\theta) = e^{-\int_0^t r(s) ds} \mathbb{E}(-r(t)X_t(\theta) + \mu(X_t(\theta), t; \theta)) dt = 0$$

Therefore, $\mathbb{E}Y_t(\theta) = x_0$ and $\mathbb{E}X_t(\theta) = B(t)x_0$. ■

Proof of Theorem 2.4: Our proof runs in parallel to that in Theorem 1.1 of [Friedman (1975)].

However, we make many adjustments to accommodate the unbounded parametrization.

First we prove the existence of the solution. Let N and Δ be positive numbers. Define

$$\phi_N(V) = \begin{cases} 1, & (N-1)\Delta \leq f(V) < N\Delta \\ 0, & \text{Otherwise.} \end{cases}$$

We define $Y_t^{(k,N)}$ by iteration as follows:

$$(7.3) \quad \begin{aligned} Y_t^{(0,N)} &= x_0, \\ Y_t^{(k+1,N)} &= x_0 + \int_0^t \phi_N(V) \mu(Y_s^{(k,N)}, s; V) ds + \int_0^t \phi_N(V) \sigma(Y_s^{(k,N)}, s; V) dW_s. \end{aligned}$$

We have the following lemmas regarding the upper bounds of $\left| Y_t^{(k+1,N)} - Y_t^{(k,N)} \right|^2$

Lemma 7.1 Define $Y_t^{(k,N)}$ as in (7.3), then

$$(7.4) \quad \mathbb{E} \left| Y_t^{(k+1,N)} - Y_t^{(k,N)} \right|^2 \leq \frac{(Mt(N\Delta)^2)^{k+1}}{(k+1)!} (1 + |x_0|)^2 \mathbb{E}(\phi_N(V))$$

where $M = 2T_0 + 2$.

Lemma 7.2 Define $Y_t^{(k,N)}$ as in (7.3), then

$$(7.5) \quad \mathbb{E} \sup_{0 \leq t \leq T_0} \left| Y_t^{(k+1,N)} - Y_t^{(k,N)} \right|^2 \leq (2T_0 + 8)(N\Delta)^2 \int_0^{T_0} \mathbb{E} \left| Y_t^{(k,N)} - Y_t^{(k-1,N)} \right|^2 ds$$

Combining (7.4) and (7.5), we have

$$\mathbb{E} \sup_{0 \leq t \leq T_0} \left| Y_t^{(k+1,N)} - Y_t^{(k,N)} \right|^2 \leq H \frac{(M(N\Delta)^2)^{k+1}}{k!} \mathbb{E}(\phi_N(V))$$

where $H = 4T_0(1 + |x_0|)^2$.

Now we define $X_s^{(k)}$ iteratively as

$$(7.6) \quad \begin{aligned} X_t^{(0)} &= x_0 \\ X_t^{(k+1)} &= x_0 + \int_0^t \mu(X_s^{(k)}, s; V) ds + \int_0^t \sigma(X_s^{(k)}, s; V) dW_s \end{aligned}$$

It is clear that $X_t^{(k)} = Y_t^{(k,N)}$, if $(N-1)\Delta \leq f(V) < N\Delta$. Then we have

$$\begin{aligned} \mathbb{E} \sup_{0 \leq t \leq T_0} |X_t^{(k+1)} - X_t^{(k)}|^2 &\leq \sum_{N=1}^{\infty} \mathbb{E} \left[\sup_{0 \leq t \leq T_0} |Y_t^{(k+1,N)} - Y_t^{(k,N)}|^2 \right] \\ &\leq H \sum_{N=1}^{\infty} \int_{(N-1)\Delta \leq f(\theta) < N\Delta} \frac{(M(N\Delta)^2)^{k+1}}{k!} m(\theta) d\theta \\ &\leq H \int_{\Theta} \frac{[M(f(\theta) + \Delta)^2]^{k+1}}{k!} m(\theta) d\theta \end{aligned}$$

Note that Δ is arbitrary. Let $\Delta \rightarrow 0$, by Dominant Convergence Theorem, we have,

$$(7.7) \quad \mathbb{E} \sup_{0 \leq t \leq T_0} |X_t^{(k+1)} - X_t^{(k)}|^2 \leq H \int_{\Theta} \frac{[Mf^2(\theta)]^{k+1}}{k!} m(\theta) d\theta$$

Hence

$$\mathbb{P} \left\{ \mathbb{E} \sup_{0 \leq t \leq T_0} |X_t^{(k+1)} - X_t^{(k)}| > \frac{1}{2^k} \right\} \leq 2^{2k} H \int_{\Theta} \frac{[Mf^2(\theta)]^{k+1}}{k!} m(\theta) d\theta$$

Because

$$\begin{aligned} \sum_{k=0}^{\infty} 2^{2k} \int_{\Theta} \frac{[Mf^2(\theta)]^{k+1}}{k!} m(\theta) d\theta &= H \int_{\Theta} Mf^2(\theta) e^{4Mf^2(\theta)} m(\theta) d\theta \\ &\leq H \int_{\Theta} e^{5Mf^2(\theta)} m(\theta) d\theta < \infty \end{aligned}$$

The Borel-Cantelli Lemma implies

$$\mathbb{P} \left\{ \mathbb{E} \sup_{0 \leq t \leq T_0} |X_t^{(k+1)} - X_t^{(k)}| > \frac{1}{2^k} \text{ i.o. } \right\} = 0$$

Then it follows that $X_0 + \sum_{i=1}^{k-1} X_t^{(i+1)} - X_t^{(i)} = X_t^{(k)}$ converge uniformly in $t \in [0, T_0]$. De-

note the limit as \tilde{X}_t . Based on the standard arguments from Theorem 1.1 of [Friedman (1975)],

we can show \tilde{X}_t satisfies the equation

$$\tilde{X}_t = x_0 + \int_0^t \mu(\tilde{X}_s, s; V) ds + \int_0^t \sigma(\tilde{X}_s, s; V) dW_s.$$

Next we show \tilde{X}_t is square integrable. From (7.7),

$$\begin{aligned}\mathbb{E} \left| X_t^{(k+1)} \right|^2 &\leq |x_0|^2 + \sum_{i=1}^k \mathbb{E} \sup_{0 \leq t \leq T_0} \left| X_t^{(i+1)} - X_t^{(i)} \right|^2 \\ &\leq H \sum_{i=0}^k \int_{\Theta} \frac{[M f^2(\theta)]^{i+1}}{i!} m(\theta) d\theta \\ &\leq H \int_{\Theta} M f^2(\theta) e^{M f^2(\theta)} m(\theta) d\theta\end{aligned}$$

Let $k \rightarrow \infty$, from Fatou's lemma, we have

$$(7.8) \quad \mathbb{E} \left| \tilde{X}_t \right|^2 < H \int_{\Theta} M f^2(\theta) e^{M f^2(\theta)} m(\theta) d\theta < \infty.$$

Therefore \tilde{X}_t is a strong solution for the SDE.

To prove the uniqueness of the solution, we assume $\tilde{X}_t^{(1)}$ and $\tilde{X}_t^{(2)}$ are two solutions. Let $\psi(V) = 1$, if $f(V) \leq N$; 0, otherwise. Note that $0 \leq \psi(V) f^2(V) \leq \psi(V) N^2$. Taking the expectation of $\psi(V) \left| \tilde{X}_t^{(1)} - \tilde{X}_t^{(2)} \right|^2$, we have

$$\begin{aligned}\mathbb{E} \left(\psi(V) \left| \tilde{X}_t^{(1)} - \tilde{X}_t^{(2)} \right|^2 \right) &\leq 2\mathbb{E} \left(\psi(V) \left| \int_0^t \left(\mu(\tilde{X}_s^{(1)}, s; V) - \mu(\tilde{X}_s^{(2)}, s; V) \right) ds \right|^2 \right) \\ &\quad + 2\mathbb{E} \left(\psi(V) \left| \int_0^t \left(\sigma(\tilde{X}_s^{(1)}, s; V) - \sigma(\tilde{X}_s^{(2)}, s; V) \right) dW_s \right|^2 \right) \\ &= 2\mathbb{E} \left| \int_0^t \psi(V) \left(\mu(\tilde{X}_s^{(1)}, s; V) - \mu(\tilde{X}_s^{(2)}, s; V) \right) ds \right|^2 \\ &\quad + 2 \int_0^t \mathbb{E} \left(\psi(V) \left| \sigma(\tilde{X}_s^{(1)}, s; V) - \sigma(\tilde{X}_s^{(2)}, s; V) \right|^2 \right) ds \\ &\leq 2t \int_0^t \mathbb{E} \left(\psi(V) f^2(V) \left| \tilde{X}_s^{(1)} - \tilde{X}_s^{(2)} \right|^2 \right) ds \\ &\quad + 2 \int_0^t \mathbb{E} \left(\psi(V) f^2(V) \left| \tilde{X}_s^{(1)} - \tilde{X}_s^{(2)} \right|^2 \right) ds \\ &\leq 2(t+1)N^2 \int_0^t \mathbb{E} \left(\psi(V) \left| \tilde{X}_s^{(1)} - \tilde{X}_s^{(2)} \right|^2 \right) ds\end{aligned}$$

Let $g(t) = \mathbb{E} \left(\psi(V) \left| \tilde{X}_t^{(1)} - \tilde{X}_t^{(2)} \right|^2 \right)$, then $g(t)$ satisfies,

$$0 \leq g(t) \leq K \int_0^t g(s) ds, \quad g(0) = 0$$

where $K = 2(T_0 + 1)N^2$ is a constant. Therefore $g(t) \equiv 0$. Let $N \rightarrow \infty$, by Dominant Convergence Theorem, we prove $\mathbb{E} \left| \tilde{X}_s^{(1)} - \tilde{X}_s^{(2)} \right|^2 ds = 0$, hence the uniqueness of the solution. ■

Proof of Lemma 7.1: When $k = 0$,

$$\left| Y_t^{(1,N)} - Y_t^{(0,N)} \right|^2 \leq 2 \left| \int_0^t \phi_N(V) \mu(x_0, s; V) ds \right|^2 + 2 \left| \int_0^t \phi_N(V) \sigma(x_0, s; V) dW_s \right|^2$$

Note that $0 \leq \phi_N(V)f(V) \leq N\Delta$. Applying the linear growth condition, we have

$$\begin{aligned} \mathbb{E} \left| Y_t^{(1,N)} - Y_t^{(0,N)} \right|^2 &\leq 2\mathbb{E} \left| \int_0^t \phi_N(V) \mu(x_0, s; V) ds \right|^2 + 2 \int_0^t \mathbb{E} \left| \phi_N(V) \sigma(x_0, s; V) \right|^2 ds \\ &\leq 2\mathbb{E} \left(\left| \int_0^t \phi_N(V) f(V) (1 + |x_0|) ds \right|^2 \right) + 2\mathbb{E} \left(\int_0^t \phi_N(V) f(V) (1 + |x_0|)^2 ds \right) \\ &\leq 2\mathbb{E} (\phi_N(V)) \left| \int_0^t (1 + |x_0|) N\Delta ds \right|^2 + 2\mathbb{E} (\phi_N(V)) \int_0^t \left((1 + |x_0|) N\Delta \right)^2 ds \\ &= \mathbb{E} (\phi_N(V)) (1 + |x_0|)^2 M t (N\Delta)^2 \end{aligned}$$

where $M = 2T_0 + 2$. Now assume (7.4) holds true for $k = 0, 1, \dots, m-1$. When $k = m$, note that

$$\begin{aligned} \left| Y_t^{(m+1,N)} - Y_t^{(m,N)} \right|^2 &\leq 2 \left| \int_0^t \phi_N(V) [\mu(Y_s^{(m,N)}, s; V) - \mu(Y_s^{(m-1,N)}, s; V)] ds \right|^2 \\ (7.9) \quad &+ 2 \left| \int_0^t \phi_N(V) [\sigma(Y_s^{(m,N)}, s; V) - \sigma(Y_s^{(m-1,N)}, s; V)] dW_s \right|^2 \end{aligned}$$

From the Lipschitz condition, we have

$$\begin{aligned} &\mathbb{E} \left| Y_t^{(m+1,N)} - Y_t^{(m,N)} \right|^2 \\ &\leq 2\mathbb{E} \left| \int_0^t \phi_N(V) f(V) |Y_s^{(m,N)} - Y_s^{(m-1,N)}| ds \right|^2 \\ &\quad + 2\mathbb{E} \int_0^t \phi_N(V) |\sigma(Y_s^{(m,N)}, s; V) - \sigma(Y_s^{(m-1,N)}, s; V)|^2 ds \\ &\leq 2(N\Delta)^2 t \mathbb{E} \int_0^t |Y_s^{(m,N)} - Y_s^{(m-1,N)}|^2 ds + 2(N\Delta)^2 \mathbb{E} \int_0^t |Y_s^{(m,N)} - Y_s^{(m-1,N)}|^2 ds \\ &\leq M(N\Delta)^2 \int_0^t \mathbb{E} |Y_s^{(m,N)} - Y_s^{(m-1,N)}| ds \end{aligned}$$

Now plugging (7.4) into the right side of the inequality, we have

$$\begin{aligned}\mathbb{E} \left| Y_t^{(m+1,N)} - Y_t^{(m,N)} \right|^2 &\leq M(N\Delta)^2 \int_0^t \frac{(Ms(N\Delta)^2)^m}{(m)!} (1 + |x_0|)^2 \mathbb{E}(\phi_N(V)) ds \\ &= \frac{(Mt(N\Delta)^2)^{m+1}}{(m+1)!} (1 + |x_0|)^2 \mathbb{E}(\phi_N(V))\end{aligned}$$

By induction, we complete our proof. ■

Proof of Lemma 7.2: From (7.9), we have

$$\begin{aligned}\sup_{0 \leq t \leq T_0} \left| Y_t^{(m+1,N)} - Y_t^{(m,N)} \right|^2 &\leq 2T_0(N\Delta)^2 \int_0^{T_0} \left| Y_s^{(m,N)} - Y_s^{(m-1,N)} \right|^2 ds \\ &\quad + 2 \sup_{0 \leq t \leq T_0} \left| \int_0^t \phi_N(V) [\sigma(Y_s^{(m,N)}, s; V) - \sigma(Y_s^{(m-1,N)}, s; V)] dW_s \right|^2\end{aligned}$$

Theorem 4.3.6 of [Friedman (1975)] shows

$$\begin{aligned}&\mathbb{E} \left\{ \sup_{0 \leq t \leq T_0} \left| \int_0^t \phi_N(V) [\sigma(Y_s^{(m,N)}, s; V) - \sigma(Y_s^{(m-1,N)}, s; V)] dW_s \right|^2 \right\} \\ &\leq 4\mathbb{E} \int_0^t \phi_N(V) \left| \sigma(Y_s^{(m,N)}, s; V) - \sigma(Y_s^{(m-1,N)}, s; V) \right|^2 dt \\ &\leq 4(N\Delta)^2 \int_0^{T_0} \mathbb{E} \left| Y_s^{(m,N)} - Y_s^{(m-1,N)} \right|^2 ds\end{aligned}$$

Therefore

$$\mathbb{E} \sup_{0 \leq t \leq T_0} \left| Y_t^{(m+1,N)} - Y_t^{(m,N)} \right|^2 \leq (2T_0 + 8)(N\Delta)^2 \int_0^{T_0} \mathbb{E} \left| Y_t^{(m,N)} - Y_t^{(m-1,N)} \right|^2 ds$$

■

Proof of Theorem 2.5: Since V is adapted to \mathcal{F}_0 , we have

$$\mathbb{E}[C(\tilde{X}.)|\mathcal{F}_0] = \mathbb{E}[C(X.(θ))]\Big|_{\theta=V}$$

Therefore,

$$\mathbb{E}[C(\tilde{X}.)] = \mathbb{E}(\mathbb{E}[C(\tilde{X}.)|\mathcal{F}_0]) = \int_{\Theta} m(\theta) \mathbb{E}[C(X.(θ))] d\theta$$

■

Proof of Theorem 2.6: Below is a list of conditions used to prove the theorem.

Let $\tilde{\sigma}(x, t) = \sqrt{\tilde{\nu}(x, t)}$. For every $x \in (a_0, \infty)$ and $t \in (0, T_0]$, we assume

A1. $m(\theta)P(x, t; \theta)$, $m(\theta)\mu(x, t; \theta)P(x, t; \theta)$, $m(\theta)\nu(x, t; \theta)P(x, t; \theta)$, and $\frac{\partial}{\partial t}m_t(\theta) \int_x^\infty P(y, t; \theta) dy$ are integrable.

A2. $\int_{\Theta} m(\theta)P(x, t; \theta) d\theta > 0$ and $\int_{\Theta} m(\theta)\nu(x, t; \theta)P(x, t; \theta) d\theta \geq 0$.

A3. There exists an integrable function $G : \Theta \rightarrow [0, \infty)$ such that for some small numbers Δx and Δt and for every $y \in [x - \Delta x, x + \Delta x]$ and $s \in [t - \Delta t, t + \Delta t]$,

$$\begin{aligned} & \left| m(\theta) \frac{\partial}{\partial y} \left(P(y, s; \theta) \mu(y, s; \theta) \right) \right| + \left| m(\theta) \frac{\partial^2}{\partial y^2} \left(P(y, s; \theta) \nu(y, s; \theta) \right) \right| \\ & + \left| m(\theta) \frac{\partial}{\partial s} P(y, s; \theta) \right| + \left| \frac{\partial}{\partial s} m_s(\theta) P(y, s; \theta) \right| \leq G(\theta) \end{aligned}$$

A4. $\tilde{\alpha}(x, t)$, $\tilde{\mu}(x, t)$ and $\tilde{\nu}(x, t)$ are locally Lipschitz, i.e., for every integer $n > 1$, there exists a positive constant C_n such that for every $0 \leq t \leq T_0$, $|x| \leq n$ and $|y| \leq n$,

$$|\tilde{\mu}(x, t) - \tilde{\mu}(y, t)| + |\tilde{\sigma}(x, t) - \tilde{\sigma}(y, t)| + |\tilde{\alpha}(x, t) - \tilde{\alpha}(y, t)| \leq C_n |x - y|$$

A5. $\tilde{\alpha}(x, t)$, $\tilde{\mu}(x, t)$ and $\tilde{\nu}(x, t)$ have linear growth, i.e., for every $0 \leq t \leq T_0$, $x \in \mathbb{R}$ and a positive constant C_0 ,

$$|\tilde{\alpha}(x, t)| + |\tilde{\mu}(x, t)| + |\tilde{\sigma}(x, t)| \leq C_0(1 + |x|)$$

A6. $P(x, t; \theta)$ has sufficient differentiability with respect to both x and t , and it satisfies the Fokker-Planck equation

$$\frac{\partial}{\partial t} P(x, t; \theta) = \frac{\partial}{\partial x} \left(P(x, t; \theta) \mu(x, t; \theta) \right) + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left(P(x, t; \theta) \nu(x, t; \theta) \right).$$

A1 and A2 ensure that $\tilde{\mu}(x, t)$, $\tilde{\nu}(x, t)$ and $\tilde{P}(x, t)$ are all well-defined. Theorem 2.2 of [Friedman (1975)] shows the unique strong solution exists under the assumption A4 and A5. We can verify the Fokker-Planck equation for the mixture distribution with

assumption A3, A6 and Dominant Convergence Theorem

$$\begin{aligned}
& \frac{\partial}{\partial t} \tilde{P}(x, t) \\
&= \int_{\Theta} \frac{\partial}{\partial t} m_t(\theta) P(x, t; \theta) d\theta + \int_{\Theta} m_t(\theta) \frac{\partial}{\partial t} P(x, t; \theta) d\theta \\
&= \int_{\Theta} \frac{\partial}{\partial t} m_t(\theta) P(x, t; \theta) d\theta - \frac{\partial}{\partial x} \left(\tilde{\mu}(x, t) \tilde{P}(x, t) \right) + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left(\tilde{\nu}(x, t) \tilde{P}(x, t) \right) \\
&= -\frac{\partial}{\partial x} \int_{\Theta} \frac{\partial}{\partial t} m_t(\theta) \int_x^{\infty} P(y, t; \theta) dy d\theta - \frac{\partial}{\partial x} \left(\tilde{\mu}(x, t) \tilde{P}(x, t) \right) + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left(\tilde{\nu}(x, t) \tilde{P}(x, t) \right) \\
&= -\frac{\partial}{\partial x} \left((\tilde{\mu}(x, t) + \tilde{\alpha}(x, t)) \tilde{P}(x, t) \right) + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left(\tilde{\nu}(x, t) \tilde{P}(x, t) \right)
\end{aligned}$$

The Fokker-Planck equation implies that $\tilde{P}(x, t)$ is the probability density function of \tilde{X}_t . ■

Proof of Proposition 2.7: Note that

$$\frac{1}{2} \frac{\partial^2}{\partial x^2} \left((x^2 \beta(x, t) - \tilde{\nu}(x, t)) \tilde{P}(x, t) \right) = \frac{\partial}{\partial x} \left((r(t)x - \tilde{\mu}(x, t) - \tilde{\alpha}(x, t)) \tilde{P}(x, t) \right)$$

Therefore $\tilde{P}(x, t)$ satisfies the Fokker-Planck equation for (2.18). ■

Proof of Corollary 2.8: Combining the linear growth condition (2.20) with the definition (2.17), we have

$$(7.10) \quad \left| \int_0^x \tilde{P}(x, t) (r(t)x - \tilde{\mu}(x, t) - \tilde{\alpha}(x, t)) dx \right| \leq K(1 + x^2) \tilde{P}(x, t).$$

Because $\tilde{P}(\cdot, t)$ has a finite second moment, based on the same argument as that in Corollary 2.3, we can prove

$$(7.11) \quad \int_0^{\infty} \tilde{P}(x, t) (r(t)x - \tilde{\mu}(x, t) - \tilde{\alpha}(x, t)) dx = 0.$$

Equivalently, we can write the equation (7.11) as

$$(7.12) \quad r(t) \int_{\Theta} m_t(\theta) \mathbb{E} X_t(\theta) d\theta - \int_{\Theta} m_t(\theta) \mathbb{E} \mu(X_t(\theta), t; \theta) d\theta - \int_{\Theta} \frac{\partial}{\partial t} m_t(\theta) \mathbb{E} X_t(\theta) d\theta = 0.$$

Let $Y_t(\theta) = m_t(\theta) e^{-\int_0^t r(s) ds} X_t(\theta)$. Then,

$$\begin{aligned}
dY_t(\theta) &= e^{-\int_0^t r(s) ds} \left(-r(t) m_t(\theta) X_t(\theta) + m_t(\theta) \mu(X_t(\theta), t; \theta) + \frac{\partial}{\partial t} m_t(\theta) X_t(\theta) \right) dt \\
&\quad + m_t(\theta) e^{-\int_0^t r(s) ds} \sqrt{\nu(X_t(\theta), t; \theta)} dW_t
\end{aligned}$$

Taking the expectation on both sides of the equation, we have

$$d\mathbb{E}Y_t(\theta) = e^{-\int_0^t r(s) ds} \mathbb{E} \left(-r(t)m_t(\theta)X_t(\theta) + m_t(\theta)\mu(X_t(\theta), t; \theta) + \frac{\partial}{\partial t}m_t(\theta)X_t(\theta) \right) dt$$

Integral both sides of the equation over $\theta \in \Theta$, then (7.12) yields

$$d \left(\int_{\Theta} \mathbb{E}Y_t(\theta) d\theta \right) = 0.$$

Therefore, $\int_{\Theta} \mathbb{E}Y_t(\theta) d\theta = x_0$ and equivalently we prove $\int_{\Theta} m_t(\theta)\mathbb{E}X_t(\theta) d\theta = B(t)x_0$. ■

Proof of Lemma 3.1: Note that the Fourier transform of normal distribution with mean μ and variance σ^2 is $\exp(\mu\eta i - \sigma^2\eta^2/2)$. This directly leads to the equation (3.4).

We know that the expected value of the random variable from the risk neutral distribution at time T is $F(T)$. On the other hand, the expect value from the lognormal mixture distribution is

$$\begin{aligned} & \int_0^\infty x \tilde{L}_T(x) dx \\ &= \int_0^\infty x \left(\int_{\Theta} m(\theta) L_T(x; \theta) d\theta \right) dx \\ &= F(T) \int_{\Theta} \left(\int_{-\infty}^\infty \frac{e^x}{\sqrt{2\pi v(\theta, T)}} \exp \left(-\frac{(x - u(\theta, T) + v(\theta, T)/2)^2}{2v(\theta, T)} \right) dx \right) m(\theta) d\theta \\ &= F(T) \int_{\Theta} m(\theta) e^{u(\theta, T)} d\theta \end{aligned}$$

This yields equation (3.5). ■

Proof of Theorem 3.2: Let $g(\theta, t) := M_t^{-1}(\hat{M}_t(\theta))$. Define

$$\hat{u}(\theta, t) := \int_0^t \hat{\mu}(\theta, s) - r(s) ds, \quad \hat{v}(\theta, t) := \int_0^t \hat{v}(\theta, s) ds.$$

It is clear that

$$\hat{u}(\theta, t) = u(g(\theta, t), t), \quad \hat{v}(\theta, t) = v(g(\theta, t), t).$$

Note that

$$(m_t(g(\theta, t)) + h(g(\theta, t))) \frac{\partial}{\partial \theta} g(\theta, t) = \hat{m}_t(\theta) + h(\theta).$$

Then we have

$$\begin{aligned}
& \int_{\Theta} \frac{m_t(\theta) + h(\theta)}{x\sqrt{2\pi v(\theta, t)}} \exp\left(-\frac{(\log(x/F(t)) - u(\theta, t) + v(\theta, t)/2)^2}{2v(\theta, t)}\right) d\theta \\
&= \int_{\Theta} \frac{m_t(g(\theta, t)) + h(g(\theta, t))}{x\sqrt{2\pi v(g(\theta, t), t)}} \exp\left(-\frac{(\log(x/F(t)) - u(g(\theta, t), t) + v(g(\theta, t), t)/2)^2}{2v(g(\theta, t), t)}\right) \frac{\partial}{\partial \theta} g(\theta, t) d\theta \\
&= \int_{\Theta} \frac{\hat{m}_t(\theta) + h(\theta)}{x\sqrt{2\pi \hat{v}(\theta, t)}} \exp\left(-\frac{(\log(x/F(t)) - \hat{u}(\theta, t) + \hat{v}(\theta, t)/2)^2}{2\hat{v}(\theta, t)}\right) d\theta
\end{aligned}$$

Therefore,

$$\begin{aligned}
& \int_{\Theta} \frac{m_t(\theta)}{x\sqrt{2\pi V(\theta, t)}} \exp\left(-\frac{(\log(x/F(t)) - u(\theta, t) + v(\theta, t)/2)^2}{2v(\theta, t)}\right) d\theta \\
&= \int_{\Theta} \frac{m_a(\theta) + h(\theta)}{x\sqrt{2\pi \hat{V}(\theta, t)}} \exp\left(-\frac{(\log(x/F(t)) - \hat{u}(\theta, t) + \hat{v}(\theta, t)/2)^2}{2\hat{v}(\theta, t)}\right) d\theta \\
&\quad - \int_{\Theta} \frac{h(\theta)}{x\sqrt{2\pi v(\theta, t)}} \exp\left(-\frac{(\log(x/F(t)) - u(\theta, t) + V(\theta, t)/2)^2}{2v(\theta, t)}\right) d\theta
\end{aligned}$$

Note that the right side of the above equation is the mixture distribution of the original mixture diffusion, the left side is that of the re-parametrized mixture diffusion. This completes our proof. ■

Proof of Corollary 3.3: This directly follows Theorem 3.2. ■

Proof of Corollary 3.4:

$$\begin{aligned}
\mathbb{E}C(\tilde{X}_.) &= \mathbb{E}f(\bar{\nu}(V, T)) = \int_{\hat{\Theta}} m_*(\theta) f(\bar{\nu}(\theta, T)) d\theta = \int_{\hat{\Theta}} m_*(\theta) f(M_T^{-1}(M_*(\theta))) d\theta \\
&= \int_0^1 f(M_T^{-1}(x)) dx = \int_0^\infty f(\theta) m_T(\theta) d\theta
\end{aligned}$$

■

Proof of Corollary 3.5: Let $R := \sup_{0 \leq t \leq T_0} |r(t)| dt$. We can set the dominant function of Theorem 2.4 as $f(\theta) := R + \sqrt{T_0 K \bar{\nu}(\theta, T_0)}$. From (3.12), we have

$$m_*(\theta) = m_{T_0}(\bar{\nu}(\theta, T_0)) \frac{\partial}{\partial \theta} \bar{\nu}(\theta, T_0)$$

Then the integrability condition $C2$ becomes

$$\begin{aligned}
\int_{\hat{\Theta}} e^{C_0 f^2(\theta)} m_*(\theta) d\theta &\leq \int_{\hat{\Theta}} e^{C_0 \left(R + \sqrt{T_0 K \bar{\nu}(\theta, T_0)}\right)^2} m_{T_0}(\bar{\nu}(\theta, T_0)) \frac{\partial}{\partial \theta} \bar{\nu}(\theta, T_0) d\theta \\
&= \int_{\Theta} e^{C_0 \left(R + \sqrt{T_0 K \theta}\right)^2} m_{T_0}(\theta) d\theta \\
&\leq e^{2C_0 R^2} \int_{\Theta} e^{2C_0 T_0 K \theta} m_{T_0}(\theta) d\theta \\
&< \infty
\end{aligned}$$

■

Proof of Lemma 4.1: It suffices to show

$$\mathcal{F}(E_t)(\eta) = \int_0^\infty m_t(\theta) e^{-(i\eta + \eta^2)\theta t/2} d\theta$$

By changing of the variable $\eta = \sqrt{\frac{2\xi}{t} - \frac{1}{4} - \frac{i}{2}}$, equivalently we have

$$(7.13) \quad \mathcal{F}(E_t) \left(\sqrt{\frac{2\eta}{t} - \frac{1}{4} - \frac{i}{2}} \right) = \mathcal{L}(m_t)(\eta)$$

From Bernstein's Theroem (see P160 [Widder (1941)]), the equation (7.13) is equivalent to that $\mathcal{F}(E_t) \left(\sqrt{\frac{2\eta}{t} - \frac{1}{4} - \frac{i}{2}} \right)$ is completely monotonic. ■

Proof of Proposition 4.2: Itô Lemma yields

$$\log(\tilde{X}_t(\vec{\theta})/F(t)) = \sum_{k=1}^n Y_t^{(k)}(\theta^{(k)})$$

where

$$Y_t^{(k)}(\theta^{(k)}) = -\frac{1}{2} \int_0^t \nu(\theta^{(k)}, s) ds + \int_0^t \sqrt{\nu_{(k)}(\theta^{(k)}, s)} dW_s^{(k)}$$

Note that $m_t^{(k)}(x) = \check{m}_t^{(k)}(\bar{\nu}^{(k)}(x, t)) \frac{d}{dx} \bar{\nu}^{(k)}(x, t)$, we have the following equation for the marginal distribution of $Y_t^{(k)}(\theta^{(k)})$

$$\begin{aligned}
&\int_0^\infty \frac{m_t^{(k)}(\theta^{(k)})}{x \sqrt{2\pi t \bar{\nu}^{(k)}(\theta^{(k)}, t)}} \exp \left(-\frac{(\log(x) + t \bar{\nu}^{(k)}(\theta^{(k)}, t)/2)^2}{2t \bar{\nu}^{(k)}(\theta^{(k)}, t)} \right) d\theta^{(k)} \\
&= \int_0^\infty \frac{\check{m}_t^{(k)}(\theta^{(k)})}{x \sqrt{2\pi t \theta^{(k)}}} \exp \left(-\frac{(\log(x) + t \theta^{(k)}/2)^2}{2t \theta^{(k)}} \right) d\theta^{(k)}
\end{aligned}$$

Therefore, the mixture distribution of $\tilde{X}_t(\vec{\theta})$ with mixing function $\prod_{k=1}^n m_t^{(k)}(\theta^{(k)})$ equals to that of $\check{X}_t(\vec{\theta})$ with mixing function $\prod_{k=1}^n \check{m}_t^{(k)}(\theta^{(k)})$, where $\check{X}_t(\vec{\theta})$ is defined as

$$\log(\check{X}_t(\vec{\theta})/F(t)) = \sum_{k=1}^n -\frac{1}{2}t\theta^{(k)} + t\theta^{(k)}W_s^{(k)}$$

In the following, we prove that when $n = 2$, $m_t = \check{m}_1 * \check{m}_2$. The marginal distribution function of the moneyness of $\check{X}_t(\vec{\theta})$ is

$$\begin{aligned} & \mathbb{P}(\log(\check{X}_t(\vec{\theta})/F(t)) = dx) / dx \\ &= \int_{-\infty}^{\infty} \left(\int_0^{\infty} \frac{\check{m}_t^{(1)}(\theta^{(1)})}{\sqrt{2\pi\theta^{(1)}t}} \exp\left(-\frac{(y + \theta^{(1)}t/2)^2}{2\theta^{(1)}t}\right) d\theta^{(1)} \right) \times \\ & \quad \left(\int_0^{\infty} \frac{\check{m}_t^{(2)}(\theta^{(2)})}{\sqrt{2\pi\theta^{(2)}t}} \exp\left(-\frac{(x - y + \theta^{(2)}t/2)^2}{2\theta^{(2)}t}\right) d\theta^{(2)} \right) dy \\ &= \int_0^{\infty} \int_0^{\infty} \frac{\check{m}_t^{(1)}(\theta^{(1)})\check{m}_t^{(2)}(\theta^{(2)})}{\sqrt{2\pi(\theta^{(1)} + \theta^{(2)})t}} \exp\left(-\frac{(x + (\theta^{(1)} + \theta^{(2)})t/2)^2}{2(\theta^{(1)} + \theta^{(2)})t}\right) d\theta^{(1)} d\theta^{(2)} \\ &= \int_0^{\infty} \frac{1}{\sqrt{2\pi\theta t}} \exp\left(-\frac{(x + \theta t/2)^2}{2\theta t}\right) \left(\int_0^{\theta} \check{m}_t^{(1)}(\psi)\check{m}_t^{(2)}(\theta - \psi) d\psi \right) d\theta \\ &= \int_0^{\infty} \frac{(\check{m}_t^{(1)} * \check{m}_t^{(2)})(\theta)}{\sqrt{2\pi\theta t}} \exp\left(-\frac{(x + \theta t/2)^2}{2\theta t}\right) d\theta \end{aligned}$$

Then induction leads to the proof for any $n > 2$. ■

Proof of Lemma 4.3: Take Fourier transform of the normal mixture distribution of (4.3), it is equivalent to derive the necessary and sufficient condition for

$$\mathcal{F}(E_T)(\eta) = \mathcal{F}(m_T)(\eta) \int_0^{\infty} b(\psi) \exp\left(-\frac{1}{2}(i\eta + \eta^2) \int_0^T \nu(\psi, s) ds\right) d\psi$$

and equivalently

$$\mathcal{F}(m_T)(\eta) = G_T(\eta)$$

This is the direct result of Bochner's Theorem [Grafakos (2008)]. ■

Proof of Corollary 4.4: Itô's Lemma shows

$$\begin{aligned} d\log(\tilde{X}_t) &= \left(r(t) - \frac{1}{2}\nu(V, t) - \frac{1}{2}\beta(\tilde{U}_t, t) \right) dt + \sqrt{\beta(\tilde{U}_t, t)} dB_t + \sqrt{\nu(V, t)} dW_t \\ &= d\log(\tilde{U}_t) + (r(t) - \nu(V, t)/2) dt + \sqrt{\nu(V, t)} dW_t \end{aligned}$$

This leads to the solution for \tilde{X}_t ,

$$(7.14) \quad \log(\tilde{X}_t/F(t)) = \log(\tilde{U}_t) - \frac{1}{2} \int_0^t \nu(V, s) ds + \int_0^t \sqrt{\nu(V, s)} dW_s$$

Compare (7.14) to the solution for $X_t(\theta, \psi)$

$$(7.15) \quad \log(X_t(\theta, \psi)/F(t)) = \theta t - \frac{1}{2} \int_0^t \nu(\psi, s) ds + \int_0^t \sqrt{\nu(\psi, s)} dW_s$$

Note that the marginal distribution function of $\log(\tilde{U}_t)$ is $tm_t(t\theta)$. Because V, \tilde{U}_t and W_t are independent to each other, the marginal density function of $\log(\tilde{X}_t/F(t))$ is exactly $E_t(\cdot)$. ■

Proof of Proposition 4.5: Note that the marginal distribution function of the money-ness of asset price is

$$N_t(x; \theta, \psi) = \frac{1}{\sqrt{2\pi\theta t}} \exp\left(-\frac{(x + \psi\theta t/2)^2}{2\theta t}\right).$$

Then we have

$$\begin{aligned} & \int_0^\infty m(\theta, \psi) N_T(x; \theta, \psi) d\theta \\ &= \frac{1}{\sqrt{2\pi\theta T}} \int_0^\infty m(\theta, \psi) \exp\left(-\frac{(x + \psi\theta T/2)^2}{2\theta T}\right) d\theta \\ &= \frac{e^{-\psi x/2}}{\sqrt{2\pi\theta T}} \int_0^\infty \sqrt{\frac{\pi}{2\theta^3 T}} \exp\left(\frac{\theta T}{8}\right) \mathcal{L}^{-1}(H_\psi)\left(\frac{1}{2\theta T}\right) \exp\left(-\frac{x^2}{2\theta T} - \frac{\theta T}{8}\right) d\theta \\ &= e^{-\psi x/2} \int_0^\infty \mathcal{L}^{-1}(H_\psi)(t) \exp(-x^2 t) dt \\ &= e^{-\psi x/2} H_\psi(x^2), \end{aligned}$$

Therefore, the normal mixture distribution $\tilde{N}_T(x)$ is

$$\begin{aligned} \tilde{N}_T(x) &= \int_0^\infty m(\theta, +) N_T(x; \theta, +) d\theta + \int_0^\infty m(\theta, -) N_T(x; \theta, -) d\theta \\ &= e^{-x/2} H_+(x^2) + e^{x/2} H_-(x^2). \end{aligned}$$

When $x \geq 0$, the normal mixture distribution becomes

$$e^{-x/2} \frac{e^{x/2} E_T(-x) - e^{-x/2} E_T(x)}{e^x - e^{-x}} + e^{x/2} \frac{e^{x/2} E_T(x) - e^{-x/2} E_T(-x)}{e^x - e^{-x}} = E_T(x).$$

Similarly, when $x \leq 0$, the normal mixture distribution becomes

$$e^{-x/2} \frac{e^{-x/2} E_T(x) - e^{x/2} E_T(-x)}{e^{-x} - e^x} + e^{x/2} \frac{e^{-x/2} E_T(-x) - e^{x/2} E_T(x)}{e^{-x} - e^x} = E_T(x).$$

Therefore, we have proved $\tilde{N}_T(x) = E_T(x)$. ■

Proof of Proposition 4.6: We have,

$$\mathcal{F}(\tilde{N}_T)(\eta) = \int_{\mathbb{R}} m(\theta) \exp\left(- (i\eta + \eta^2)\nu_0 T/2\right) e^{iT\eta\theta} d\theta = \mathcal{F}(E_T)(\eta)$$

Since E_T is continuous and $\mathcal{F}(E_T)(\eta)$ is invertible, we have $\tilde{N}_T(x) = E_T(x)$. ■

Proof of Corollary 4.7: Define the normal distribution function

$$\phi(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\mu^2/(2\sigma^2)).$$

Note that the normal density function of (3.1) can be expressed as the convolution of two normal distributions,

$$\begin{aligned} & \phi(x - U(\theta, T) + V(\theta, T)/2, V(\theta, T)) \\ &= T \int_{-\infty}^{\infty} \phi(x - T\psi + TV_0/2, TV_0) \phi(T\psi - U(\theta, T) + (V(\theta, T) - TV_0)/2, V(\theta, T) - TV_0) d\psi \end{aligned}$$

The mixture distribution of $X_t(\theta)$ in (3.1) then becomes

$$\begin{aligned} & \int_{\Theta} m(\theta) N_T(x; \theta) d\theta \\ &= \int_{-\infty}^{\infty} \phi(x - T\psi + TV_0/2, V_0) \int_{\Theta} m(\theta) \phi(T\psi - U(\theta, t) + (V(\theta, T) - TV_0)/2, V(\theta, t) - TV_0) d\theta d\psi \end{aligned}$$

Therefore, the mixture distribution of (3.1) is identical to that of (4.8) with the mixing function $\hat{m}(\theta)$ and the instantaneous variance V_0 , where

$$\hat{m}(\psi) = \int_{\Theta} m(\theta) \phi(\psi - U(\theta, T) + (V(\theta, T) - TV_0)/2, V(\theta, T) - TV_0) d\theta.$$

■

Proof of Proposition 4.8: The proof is similar to Proposition 4.6, we omit the details. ■

Proof of Theorem 5.2: First, we prove the zero-coupon bond process $\hat{B}(t, T) := e^{-\int_t^T \hat{f}(t, u) du}$ is risk-neutral. Let $\hat{Y}(t, T) := \int_t^T \hat{f}(t, u) du$. Note that $\hat{\phi}^{(k)}(t, t) = 0$. Leibniz rule yields

$$\begin{aligned} d\hat{Y}(t, T) &= -\hat{f}(t, t) dt + \int_t^T d\hat{f}(t, u) du \\ &= -\hat{f}(t, t) dt + \frac{1}{2} \sum_{k=1}^n (\hat{\phi}^{(k)}(t, T))^2 \hat{\nu}^{(k)}(\hat{V}^{(k)}, t) dt + \sum_{k=1}^n \hat{\phi}^{(k)}(t, T) \sqrt{\hat{\nu}^{(k)}(\hat{V}^{(k)}, t)} dW_t^{(k)} \end{aligned}$$

Note that $\hat{B}(t, T) = e^{-\hat{Y}(t, T)}$. Itô's Lemma yields

$$\frac{d\hat{B}(t, T)}{\hat{B}(t, T)} = \hat{f}(t, t) dt - \sum_{k=1}^n \hat{\phi}^{(k)}(t, T) \sqrt{\hat{\nu}^{(k)}(\hat{V}^{(k)}, t)} dW_t^{(k)}$$

Since $r(t) := \hat{f}(t, t)$ is the spot risk-free interest rate, $\hat{B}(t, T)$ is a risk-neutral process with numeraire $e^{\int_0^t r(s) ds}$.

Next we prove the forward LIBOR process $\hat{L}(t, T_{(k-1)}, T_{(k)})$ has the marginal density distribution $D_t^{(k)}(\cdot)$. Let $\hat{\tau}(t) = \min_{T_{(k)} \geq t} k$. When $m > \hat{\tau}(t)$, we have $\hat{\phi}^{(k)}(t, T_{(m)}) = 1$. Therefore,

$$\begin{aligned} d\hat{Y}(t, T_{(m)}) &= -\hat{f}(t, t) dt + \frac{1}{2} \sum_{k=\hat{\tau}(t)+1}^m \hat{\nu}^{(k)}(\hat{V}^{(k)}, t) dt + \sum_{k=\hat{\tau}(t)+1}^m \sqrt{\hat{\nu}^{(k)}(\hat{V}^{(k)}, t)} dW_t^{(k)} \\ &\quad + \frac{1}{2} \sum_{k=1}^{\hat{\tau}(t)} (\hat{\phi}^{(k)}(t, T_{(m)}))^2 \hat{\nu}^{(k)}(\hat{V}^{(k)}, t) dt + \sum_{k=1}^{\hat{\tau}(t)} \hat{\phi}^{(k)}(t, T_{(m)}) \sqrt{\hat{\nu}^{(k)}(\hat{V}^{(k)}, t)} dW_t^{(k)} \end{aligned}$$

Let $\hat{Y}(t, T_1, T_2) = \hat{Y}(t, T_2) - \hat{Y}(t, T_1)$. When $t \leq T_{(k-1)}$, we have $\hat{\tau}(t) \leq k-1$ and

$$d\hat{Y}(t, T_{(k-1)}, T_{(k)}) = \frac{1}{2} \hat{\nu}^{(k)}(\hat{V}^{(k)}, t) dt + \sqrt{\hat{\nu}^{(k)}(\hat{V}^{(k)}, t)} dW_t^{(k)}$$

By the definition of the solution of mixture diffusion, $\hat{Y}(t, T_{(k-1)}, T_{(k)})$ has the mixture distribution $E_t^{(k)}(\cdot)$ for every $0 < t \leq T_{(k-1)}$. Consequently, the forward LIBOR process,

$$\hat{L}(t, T_{(k-1)}, T_{(k)}) = \frac{\hat{B}(t, T_{(k-1)})/\hat{B}(t, T_{(k)}) - 1}{T_{(k)} - T_{(k-1)}} = \frac{e^{-\hat{Y}(t, T_{(k-1)}, T_{(k)})} - 1}{T_{(k)} - T_{(k-1)}}$$

has the marginal density distribution $D_t^{(k)}(\cdot)$ for every $0 < t \leq T_{(k-1)}$. ■

Proof of Proposition 5.3: First, we observe that when $t \leq T_{(\alpha)}$ and $m, k \geq \alpha + 1$, $\hat{\phi}^{(k)}(t, T_{(m)}) = 1$ if $m \geq k$; 0, otherwise. Therefore, we have the following simplified formula for zero-coupon bond process

$$\hat{B}(t, T_{(m)}) = \hat{B}(t, T_{(\alpha)}) \exp \left(\sum_{k=\alpha+1}^m -\frac{1}{2} \int_0^t \hat{\nu}^{(k)}(\hat{V}^{(k)}, s) ds + \int_0^t \sqrt{\hat{\nu}^{(k)}(\hat{V}^{(k)}, s)} dW_s^{(k)} \right)$$

We have

$$\begin{aligned} & \mathbb{E} \left[\sum_{m=\alpha}^{\beta} c_{(m)}(K) B^{-1}(t) \hat{B}(t, T_{(m)}) \right]^+ \\ &= \mathbb{E} \left[B^{-1}(t) \hat{B}(t, T_{(\alpha)}) \left(1 + \sum_{m=\alpha+1}^{\beta} c_{(m)}(K) \times \right. \right. \\ & \quad \left. \left. \exp \left(\sum_{k=\alpha+1}^m -\frac{1}{2} \int_0^t \hat{\nu}^{(k)}(\hat{V}^{(k)}, s) ds + \int_0^t \sqrt{\hat{\nu}^{(k)}(\hat{V}^{(k)}, s)} dW_s^{(k)} \right) \right) \right]^+ \\ &= \mathbb{E} \left[1 + \sum_{m=\alpha+1}^{\beta} c_{(m)}(K) \exp \left(\sum_{k=\alpha+1}^m \int_0^t \hat{\nu}^{(k)}(\hat{V}^{(k)}, s) ds + \int_0^t \sqrt{\hat{\nu}^{(k)}(\hat{V}^{(k)}, s)} dW_s^{(k)} \right) \right]^+ \end{aligned}$$

The last equation is because conditionally on \mathcal{F}_0 , $B^{-1}(t) \hat{B}(t, T_{(\alpha)})$ is a martingale and is independent of $W_t^{(\alpha+1)}, \dots, W_t^{(\beta)}$. Let

$$N^{(k)} := \int_0^t \sqrt{\hat{\nu}^{(k)}(\hat{V}^{(k)}, s)} dW_s^{(k)} \bigg/ \sqrt{\int_0^t \hat{\nu}^{(k)}(x, s) ds}$$

Then $N^{(k)}$ are i.i.d. standard normal random variables and they are independent of \hat{V} .

This completes our proof. ■

Proof of Proposition 5.4: We skip the proof that $\tilde{B}(t, T)$ is a risk-neutral process because it is the result of the same logic as that in Theorem 5.2.

First we prove the forward LIBOR process $\hat{L}(t, T_{(k-1)}, T_{(k)})$ still has the marginal density distribution $D_t^{(k)}(\cdot)$. Let $\tilde{Y}(t, T) := \int_t^T \tilde{f}(t, u) du$. Note that $\tilde{\phi}^{(k)}(t, t) = 0$. Leibniz rule yields

$$d\tilde{Y}(t, T) = -\tilde{f}(t, t) dt + \frac{1}{2} \sum_{k=1}^n (\tilde{\phi}^{(k)}(t, T))^2 \tilde{\nu}^{(k)}(\tilde{V}^{(k)}, t) dt + \sum_{k=1}^n \tilde{\phi}^{(k)}(t, T) \sqrt{\tilde{\nu}^{(k)}(\tilde{V}^{(k)}, t)} d\tilde{W}_t^{(k)}$$

Based on our definition of $\tilde{\phi}^{(k)}(t, T)$, we have $\tilde{Y}(t, T_{(k)}) = \hat{Y}(t, T_{(k-1)})$ when $t \leq T_{(t-1)}$. Therefore, $\tilde{B}(t, T_{(k-1)})/\tilde{B}(t, T_{(k)}) = \hat{B}(t, T_{(k-1)})/\hat{B}(t, T_{(k)})$. Consequently, $\tilde{L}(t, T_{(k-1)}, T_{(k)}) = \hat{L}(t, T_{(k-1)}, T_{(k)})$ and its marginal distribution is $D_t^{(k)}(\cdot)$.

Next we prove that marginal density function of $\tilde{L}(T_{(t-1)}, T_{(t-1)}, T_{(t-.5)})$ is $D_A^{(k)}(\cdot)$, and that of $\tilde{L}(T_{(t-.5)}, T_{(t-.5)}, T_{(t)})$ is $D_B^{(k)}(\cdot)$. Let $\tilde{Y}(t, T_1, T_2) = \tilde{Y}(t, T_2) - \tilde{Y}(t, T_1)$. From the definition of $\tilde{\phi}^{(k)}(t, T)$, we have

$$\tilde{Y}(t, T_{(k-1)}, T_{(k-.5)}) = \frac{1}{2} \int_0^t \tilde{\nu}^{(k)}(\tilde{V}^{(k)}, s) ds + \int_0^t \sqrt{\tilde{\nu}^{(k)}(\tilde{V}^{(k)}, t)} d\tilde{W}_t^{(k)}$$

Based on our definition, $\tilde{Y}(T_{(t-1)}, T_{(k-1)}, T_{(k-.5)})$ has the marginal density function $\tilde{F}_A^{(k)}(x)$ and $\hat{Y}(T_{(k-1)}, T_{(k-1)}, T_{(k-.5)})$ has the marginal density function $\hat{F}_A^{(k)}(x)$. Note that the *moneyness* of $\tilde{L}(T_{(t-1)}, T_{(k-1)}, T_{(k-.5)})$ is

$$\tilde{Y}(T_{(t-1)}, T_{(k-1)}, T_{(k-.5)}) + \hat{Y}(T_{(t-1)}, T_{(k-1)}, T_{(k-.5)}).$$

Because \tilde{Y} and \hat{Y} are independent to each other, the *moneyness* of LIBOR has the marginal density function

$$\hat{F}_A^{(k)} * \tilde{F}_A^{(k)} = E_A^{(k)}.$$

Consequently, the the marginal density function of $\tilde{L}(T_{(t-1)}, T_{(t-1)}, T_{(t-.5)})$ is $D_A^{(k)}(\cdot)$.

Similarly when $t = T_{(k-.5)}$ and $T = T_{(k)}$, we have

$$\tilde{Y}(T_{(k-.5)}, T_{(k-.5)}, T_{(k)}) = -\frac{1}{2} \int_0^{T_{(k-.5)}} \tilde{\nu}^{(k)}(\tilde{V}^{(k)}, s) ds - \int_0^{T_{(k-.5)}} \sqrt{\tilde{\nu}^{(k)}(\tilde{V}^{(k)}, s)} d\tilde{W}_s^{(k)}$$

Therefore, $\tilde{Y}(T_{(k-.5)}, T_{(k)}, T_{(k-.5)})$ has the marginal density function $\tilde{F}_B^{(k)}(x)$. The same arguments as above prove that the marginal density function of $\tilde{L}(T_{(t-.5)}, T_{(t-.5)}, T_{(t)})$ is $D_B^{(k)}(\cdot)$. This completes our proof. ■

References

- [Alexander (2004)] Alexander C. (2004) Normal mixture diffusion with uncertain volatility: modelling short and long term smile effects. *Journal of Banking and Finance* 28, 2957–2980.

- [Bahra (1997)] Bahra B. (1997) Implied risk-neutral probability density functions from option prices: theory and application. *Bank of England*.
- [Brace *et al.* (1997)] Brace A., Gatarek D., Musiela M. (1997) The market model of interest rate dynamics. *Mathematical Finance* 7(2), 127-154.
- [Breedon, Litzenberger(1978)] Breedon D. T., Litzenberger, R. H. (1978) Prices of state-contingent claims implicit in option prices. *Journal of Business* 51, 621–651.
- [Brigo (2002)] Brigo D. (2002) The general mixture–diffusion SDE and its relationship with an uncertain-volatility option model with volatility-asset decorrelation. *arXiv.org*
- [Brigo *et al* (2002)] Brigo D., Mercurio F., Rapisarda F. (2002) Lognormal-mixture dynamics and calibration to market volatility smiles. *International Journal of Theoretical and Applied Finance* 5, 427–446.
- [Brigo, Mercurio (2006)] Brigo D., Mercurio F. (2006) Interest rate models – theory and practice with smile, inflation and credit. *New York: Springer-Verlag*.
- [Dupire (1994)] Dupire B. (1994) Pricing with a smile. *Risk* 7, 32–39.
- [Grafakos (2008)] Grafakos L. (2008) Classical fourier analysis. *New York: Springer-Verlag*.
- [Gulisashvili (2012)] Gulisashvili A. (2012) Analytically tractable stochastic stock price models. *New York: Springer-Verlag*.
- [Hagan *et al* (2002)] Hagan P. S., Kumar D., Lesniewski A. S., Woodward D. E. (2002) Managing smile risk. *Wilmott Magazine, September*, 1(1)
- [Heston (1993)] Heston S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies* 6(2), 327–343.

- [Friedman (1975)] Friedman A. (1975) Stochastic differential equations and applications Vol.1 *Academic Press, New York*
- [Jackwerth (2004)] Jackwerth M. (2004) Option-implied risk-neutral distributions and risk aversion. *Research Foundation of AIMR, Charlottesville, USA*.
- [Karatzas, Shreve (1988)] Karatzas I., Shreve S. E. (1988) Brownian motion and stochastic calculus. *New York: Springer-Verlag*.
- [McLachlan, Peel (2004)] McLachlan G., Peel D. (2004) Finite mixture models. *Wiley, New York*.
- [Ritchey (1990)] Ritchey R. J. (1990) Call option valuation for discrete normal mixtures. *Journal of Financial Research*, 13(4), 285–296.
- [Widder (1941)] Widder D. V. (1941) The Laplace transform. *Princeton University Press, Princeton, NJ*.